JAN 17 1923

PERIODICAL ROOM
GENERAL LIBRARY
UNIV. OF MICH.

The Journal of Educational Psychology

Devoted Primarily to the Scientific Study of Problems of Learning and Teaching.

BOARD OF EDITORS:

HAROLD ORDWAY RUGG, *Chairman.*
Lincoln School of Teachers College.
Teachers College, Columbia University.

JAMES CARLETON BELL,
Brooklyn Training School for Teachers.

FRANK NUGENT FREEMAN,
University of Chicago.

ARTHUR IRVING GATES,
Teachers College, Columbia University.

VIVIAN ALLEN CHARLES HENMON,
University of Wisconsin.

RUDOLF PINTNER,
Teachers College, Columbia University.

BEARDSLEY RUML,
Carnegie Corporation, New York City.

LEWIS MADISON TERMAN,
Leland Stanford University.

EDWARD LEE THORNDIKE,
Teachers College, Columbia University.

LAURA ZIRBES *Assistant Editor.*
Lincoln School, of Teachers College.

VOL. XIV.
No. 1.

JANUARY, 1923

$4.00 a year.
60c. per copy.

CONTENTS

MENTAL TESTS AND LINQUISTIC ABILITY. *Stephen S. Colvin and Richard D. Allen* 1

THE EFFECT OF ENCOURAGEMENT AND OF DISCOURAGEMENT UPON PERFORMANCE. *Georgina Strickland Gates and Louise Q. Rissland.* 21

WHAT SHALL WE EXPECT OF THE AQ? *Herbert A. Toops and P. W. Symonds.* 27

REVISING THE DEARBORN INTELLIGENCE EXAMINATIONS. *Walter F. Dearborn and Edward A. Lincoln* 39

NOTES ON ARTICLES IN EDUCATIONAL PSYCHOLOGY IN CURRENT ISSUES OF OTHER MAGAZINES 47

NEW PUBLICATIONS IN EDUCATIONAL PSYCHOLOGY AND RELATED FIELDS OF EDUCATION. 52

Published Monthly Except June to August by
WARWICK and YORK, Inc.,
York, Pa. *Baltimore, Md.*

ered as Second Class matter November 15, 1921, at the Post Office at York, Pennsylvania, under the Act of March 3, 1879.

The Journal of Educational Psychology

is published monthly, by Warwick & York, Inc., Baltimore, Md., and York, Pa. The issues for 1922 make Volume XIII. Title page and index are bound in December number.

Manuscripts for publication, books or other materials for review, and news items should be addressed to Harold Ordway Rugg, Chairman of the Editorial Board, 59 Edgecliffe Terrace, Yonkers, New York.

The price of the Journal is $4.00 a year, payable in advance. Foreign postage is 30 cents extra. Single issues are 60 cents each, and less than a full year's subscription is charged at single issue rate for every month ordered.

Subscribers should notify the publishers of change in address at least four weeks before publication of the issue with which change is to take effect. No claim for non-receipt of an issue will be entertained unless made within two weeks after receipt of the next succeeding issue.

Warwick & York, Inc.,
Baltimore, Md.

TEACHERS' HEADQUARTERS

European
Plan

Entirely
Fire-Proof

Centrally
Located

EDWARD DAVIS
Manager

HOTEL RENNERT
BALTIMORE, MARYLAND

University of Virginia Summer Quarter

First Term, June 18–July 28
Second Term, July 30–September 1

Courses for Elementary Teachers
Courses for High School Teachers
Courses for College Credit

The Summer Quarter is an integral part of the University Year, the courses being the same in character and credit value as in the other quarters of the year.

Degrees are conferred upon men and women for summer work.

The Master's Degree may be obtained in three Summer Quarters.

It offers opportunities unexcelled in the South and makes a strong appeal to teachers seeking broader scholarship and training, and wider social contacts, and to college students desiring to complete degree requirements.

Attendance last Quarter, 2977 from thirty-nine states and foreign countries.

The most beautiful and unique campus in America.

Accommodations at reasonable rates. Tuition for non-Virginia students $20.00 per term.

Entertainments, Music Festival, excursions.

For illustrated folder and full announcement, write to

CHAS. G. MAPHIS, Dean,
University, Virginia.

in'dis-pen'sa-ble

"Impossible to be dispensed with or done without: absolutely necessary or requisite."

This word best describes
WEBSTER'S NEW INTERNATIONAL DICTIONARY

In the schoolroom and in the library of every teacher.

WHEN QUESTIONS arise in the history recitation, in language work, in spelling, or about noted people, places, foreign words, synonyms, pronunciation, new words, flags, state seals, etc., you will find the answer in this "Supreme Authority." There is a need in *every class every day.*

WHY not suggest to your principal or superintendent that a copy be supplied for your school?

Write for Specimen Pages, Terms, etc.

2,700 Pages
6,000 Illus.
400,000 Words

G. & C. MERRIAM CO., Springfield, Mass.

THE JOURNAL OF
EDUCATIONAL PSYCHOLOGY

VolumeXIV January, 1923 Number 1

MENTAL TESTS AND LINGUISTIC ABILITY

STEPHEN S. COLVIN

Brown University

and

RICHARD D. ALLEN

Director of Research and Guidance, Public Schools, Providence

When Yerkes in 1915 published his Point Scale[1] Revision of the Binet Tests, he included among his findings data relating to the social and racial status of children tested. He concludes "that conditions which are in part describable as sociological are correlated with differences in intellectual performance, which may amount to as much as 30 per cent of the total." He says: "In view of this fact, which our results amply demonstrate, it is obviously unfair to judge by the same norm of intelligence two children, the one of whom comes from an excellent home and neighborhood and the other from a medium to poor home and neighborhood.[2]"

This caution, so definitely stated by Yerkes, has by no means always been followed and the result is that conclusions, in many cases apparently unwarranted, or at least unproven, have been drawn in regard to the levels of intelligence of various social, economic and industrial groups.

Even Yerkes himself has accepted the results of the Army Tests as definitely showing the fact that there are clear grades of intelligence among various occupational groups. He places at the top the professional group and at the bottom the unskilled laborer.[3] While it is probable that this classification has some justification, it is evident

[1] Yerkes, Robert M., Bridges, James W. and Hardwick, Rose E.: "A Point Scale for Measuring Mental Ability." Warwick and York.

[2] *Op. cit.*, p. 82.

[3] "Army Mental Tests." New York, 1920, pp. 197–200.

that the workers in various occupations represent individuals varying greatly in economic and social status.

Some eminent authorities seem to hold to the opinion that environment is of but slight importance in determining the results of mental testing—Terman[1] believes "that the environment of the home affects the results but little."[2] Even "limited acquaintance with the language employed in the examination does not put the subject to great disadvantage in many of the tests." However, if it is true, as it seems reasonable to conclude, that mental tests are to be relied upon only when those who are tested have had the same or at least similar opportunities to become acquainted with the materials of the tests and the same interests in learning about these materials, then these environmental factors are of clear significance. In linguistic tests in particular it is of large moment that those tested have a substantially similar familarity with the words employed and have a similar skill in the use of these words.

The dependence of linguistic knowledge and skill on the ability to secure average or high scores in intelligence tests of the verbal type has been pointed out by several investigators. Whipple, for example, in a summary of his results secured at the University of Michigan says[3] that 94 per cent of the students tested received in the Army Alpha grades of *B* or better, and adds: "of the remaining 6 per cent, several were students of foreign extraction whose low score must have been in a considerable measure produced by a lack of ready command of English."

Burt,[4] who has recently given the Binet-Simon scale to London school children and has worked over his findings by careful statistical methods, concludes that various factors affect the result. "Sex influences it but little; social status rather more; educational and particularly linguistic attainments more profoundly than any other factor measurable with exactitude."

In a paper read last February before the National Society for the Study of Education in its meeting at Chicago, Rugg gave an interesting summary of the correlation coefficients obtained between intelli-

---

[1] The Intelligence of School Children." Houghton, Mifflin, Boston, 1919, p. 14.

[2] *Op. cit.*, p. 12.

[3] *Twenty-first Yearbook* of the National Society for the Study of Education, p. 266. Public School Publishing Co., Bloomington, Ill.

[4] Burt, Cyril: "Mental and Scholastic Tests." Published by the London County Council, 1921, p. 208.

gence tests and school attainment. He finds in his review of various investigations that the coefficients are only moderate between scores in intelligence tests and achievement in "most non-verbal" educational tests. They are higher in the case of "somewhat verbal" educational tests and they are conspicuously high in the case of the "most verbal educational tests." The correlations between scores in verbal intelligence tests and vocabulary and reading attainments are particularly high.

Sometimes, no doubt, fluency in the use of the vernacular results in an individual receiving a higher score in an intelligence test of the linguistic type than his real mentality warrants. Downey is of the opinion that "undoubtedly the most important source of error in judging intelligence is the undue emphasis laid upon verbal fluency as a measure of social and general intelligence." This probably is a significant factor in the selection of college students for the Greek Letter fraternities. A freshman who can talk makes a decided impression on his fellow students, and in the hurry of the "rushing season" this first impression has an undue influence. Recently at Brown it has been found that the members of the freshman class pledged to fraternities received the following distribution in intelligence scores—best two-fifths, 31 per cent; lowest two-fifths, 56 per cent. Among the unpledged men, excluding those cases where because of various affiliations, etc. the men would not be "rushed" the distribution was—best two-fifths, 51 per cent; lowest two-fifths, 31 per cent. This is a striking demonstration of the intellectual superiority of the unpledged men. It is fair to assume that the pledged students in the entering class made an impression because of their verbal fluency quite out of proportion to their real ability. Apparently, too, this verbal fluency was but "skin deep." Clearly it did not serve the men who possessed it substantia ly when they were subjected to the test of the psychological examination. Goddard, like Downey, believes that high grades of mental deficiency are liable to go undetected when accompanied by verbal fluency.

However, while verbal ability may raise intelligence scores in some instances above the level of the actual intelligence of the person examined, its most marked effect is noted under the conditions when the lack of such a facility unduly lowers such scores.

The attention of one of the writers (C.) was first definitely called to this fact when in the winter of 1919–1920, the Otis Group Intelligence Scale, Forms A and B, was given under his direction by the

teachers in the elementary schools of Brookline, Massachusetts, to 1877 children in Grades IV to IX inclusive.[1]

The scores obtained were considerably above the Otis norms, particularly in the so-called "better districts" in Brookline. Otis Tests given at about the same time by Warren W. Coxe to the public school children of Cincinnati, Ohio, obtained quite different results. In this latter instance, the norms were below those published by Otis. The tests in both cities seem to have been given with due care and the disparity of results can hardly be due to differences of method or caution in administration. On the other hand, it cannot be assumed that a child in Brookline of 12 years of age has a mental age on the average 2 years in advance of a Cincinnati child of the same chronological age. The pronounced differences in scores between the Brookline children and the Cincinnati children may reasonably be attributed to differences in opportunities to learn words and acquire skill is their use. This conclusion is supported by the fact that the children in the poorer localities in Brookline did not score as high in the entire test as did the children in the more favored localities. However, in the arithmetic test (largely non-verbal in its nature) their scores were not inferior to those made by the children in the "better" localities.

Another piece of evidence in support of the above viewpoint has been found in examining the records of Brown University students in the light of their psychological tests. As a rule, men who score low in their psychological tests make a poor record in college. For example, out of a total of 95 men in last year's junior and senior classes who received low grades in their combined psychological tests, 70 have done distinctly poor work, 17 fair work, and only 8 good work. On investigating more carefully the records of these 8 men, it was found that at least 2 suffered from language handicaps, while the others were either slow thinkers, were indifferent or suffered emotional upsets. Further investigations have revealed at least 10 men in Brown University who have received psychological scores decidedly lower than their real mental ability because of language deficiencies. By *deficiencies* is not necessarily meant inability to speak English with correct pronunciation and reasonable fluency. The most conspicuous examples of language deficiencies have been found among students of Italian parentage. These students were born in America and have received their education in the public schools. They, however, have

[1] *Journal Educational Research*, Vol. III, No. 1, Jan., 1921.

lived in the Italian section and at home frequently converse in Italian and think in Italian. They have a limited English vocabulary and tend to think slowly in English.

Recently the writer (A.), Director of Research and Guidance in the public schools of Providence, has made an investigation of 50 children of American parentage and 50 of Italian parentage in the public schools of the city. The school status of these two groups of children is practically identical. They are taken from Grades V to VIII, inclusive; the ages of the American born children range from 11 to 16, those of the Italian born from 11 to 15; the average chronological age of the American born is 13.12 years, of the Italian 13.11 years; the average pedagogical age of the American children is 11.77 years, of the Italian 12. The school records of the Italian group indicate them to be in every way the equal in intelligence of the American group. When measured by the Stanford-Binet scale, the average IQ's as well as the distribution of the IQ's is only slightly in favor of the American group. However, when tested by the National Intelligence Tests, the average total score of the American children is 103; of the Italian children 90. In one test alone are the Italian children on the average equal to the American children and that is in the arithmetic test. Further, in the American group the agreement between the Binet scores and the National scores, is much closer than in the Italian group, the Italian children scoring decidedly lower in the National Tests than would be expected in the light of their school standing and of their IQ's derived from the Stanford-Binet.. The explanation of these contradictions seems obvious. The Italian children are suffering from a language handicap; hence their intelligence as determined by their scores in the verbal group Tests, is rated decidedly too low. These facts are expressed in tabular form in Table 1.

Diagram I shows in detail the relation in the two groups between the Terman and the National scores expressed in mental ages. It will be observed that while the children in both groups tend to be rated lower in mental age by the National than by the Terman Tests, the difference is considerably greater in the Italian than in the American group. It is reasonable to suppose that the individual Terman tests are a more accurate determination of intelligence than the Group National Tests; further that the linguistic factor is less important in the former than in the latter tests. It would seem evident, then, that the National Tests do not give an accurate means of determining the mental ages of the Italian pupils and that the verbal

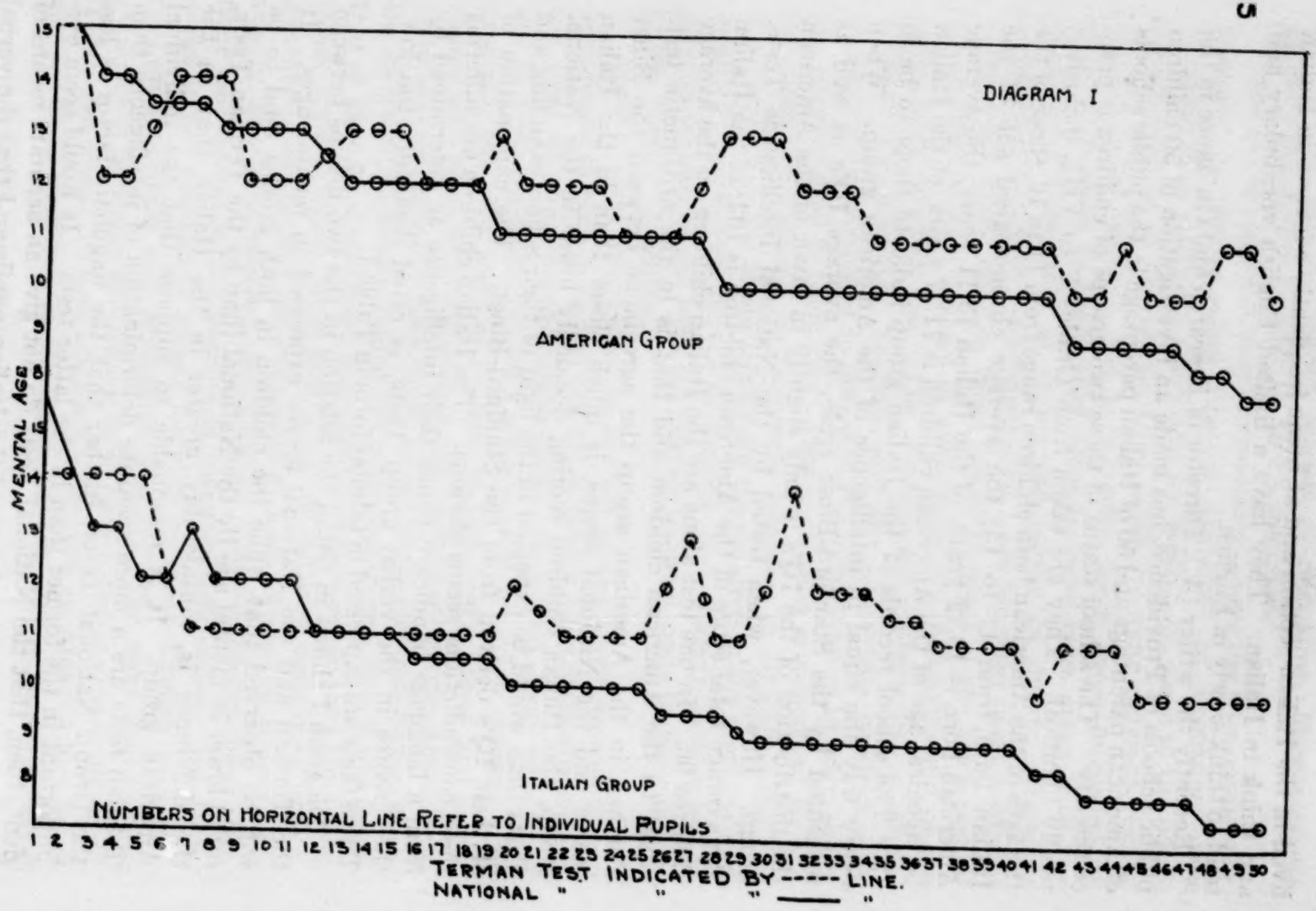
DIAGRAM I
AMERICAN GROUP
ITALIAN GROUP
MENTAL AGE
NUMBERS ON HORIZONTAL LINE REFER TO INDIVIDUAL PUPILS
TERMAN TEST INDICATED BY ----- LINE.
NATIONAL " " " ——— "

factor has contributed in lowering the individual scores. It is to be noted further from the diagram that the Terman and the National Tests for both groups are more closely in agreement at the higher mental ages than at the lower. The most striking divergence begins at about the mental age of 10, while from 11 to 15, including 19 pupils in the Italian group and 26 in the American group, the differences are unimportant.

Another comparison shows the greater agreement in the American group than in the Italian group between the mental ages derived from these two tests. By consulting Table II we find listed for the American group the mental ages of each pupil according to the National scores (N) and according to the Terman scores (T); also the difference, plus and minus, between the N mental ages and the T mental ages under column D. For the highest 25 in mental ages the total difference is 17.5, 10 plus and 7.5 minus, a very substantial agreement. For the lowest 25 the deviation totals 38, all plus.

In Table III we have a similar comparison for the Italian group. Here we find for the higher 25 in mental age a total difference of 21, with a plus difference of 14. In the lowest 25 we find no minus differences and a total plus difference of 63. Here again is a clear evidence that there is a much greater difference in the mental ages derived from the two tests in the Italian group than in the American group and also that in the Italian group the mental ages derived from the Terman scores are significantly higher than these ages derived from the National scores.

A detailed examination of Table IV, which shows the relation between the gross scores in the National Tests and grade placement for both groups, reveals the following facts: 15 Italian children received scores of 100 or above. All of these are found in the two upper grades. Twenty-three American children received similar scores, and 5 of this number are in Grade VI. Thirty-five Italian children received scores below 100 and of these 19 were in the two upper grades, while of the American children only 5 were in the two upper grades. In the two lower grades are found 16 Italian children and 22 American children. Thus judged by school attainment the Italian children are clearly less accurately classified in intelligence by the National Tests than are the American children. The objection may be raised that the school attainment of the two groups may not be a valid indication of their respective abilities, since the Americans and the Italians may be judged by different standards.

This is, however, not likely to be the case since the American and the Italian children were in the same schools and in the same rooms and in about equal proportions.

Another significant piece of evidence indicating that the Italian children in the Providence public schools suffer in their tests from a linguistic handicap is shown from comparative scores of a group of American and of Italian children in the Grade V*A*. One hundred and seventy-three American children and 163 Italian children were given the Lippincott-Chapman Classroom Products Survey Tests. These tests are composed of two sub-tests in Arithmetic (fundamentals and problems) and two reading sub-tests (selections and continuous passages). These tests, although designed to measure school achievement, are in reality for all essential purposes intelligence tests, since they are made up of elements commonly found in intelligence tests. Tables V to IX inclusive indicate the distribution of scores made by individual pupils.

In these tables the figures at the left under "score" give the standards of achievement for the various levels of the various grades according to the published norms of Chapman. In the present article the highest levels, 9 and 9+, have been combined under "nine," and the lowest levels, 1 and 1−, have been combined under "one." A study of these tables shows that the two groups are nearly equal in ability in fundamentals in arithmetic. Here language plays no definite part. In arithmetic problems the language handicap begins to be felt, as shown by the greater number of Americans above the grade median. In both of the reading tests most of the Americans are above the grade median, while most of the Italians are below. Chart IX shows how the language handicap affects the total score made by the Italians.

Still another bit of corroborative evidence as to the fact that linguistic ability plays an important rôle in tests of the verbal type is furnished by results obtained from time to time through tests given in the Mary C. Wheeler school for girls. This school is a high grade private school located on the East Side in the city of Providence and draws largely its children from families of superior social and economic status. The Otis Tests given in this school in the academic year 1919–20 "show, from the fourth to the seventh grades inclusive, that considerably more than half of the pupils tested fall in the classes near genius, very superior, or superior, while only a small percentage are below the average."[1] Later tests made

[1] *Journal of Educational Research, op. cit.*

during the academic year 1921–22 are in agreement with these earlier findings. Tests given in Grades VI and VII to 29 children show all to be above the Otis norms. Thirty-six children given the National Tests in Grades V, VI and VII have with a single exception IQ's at 100 or above, 27 being above. The median IQ is 116, the highest 150 and the lowest 96. The school achievement does not indicate that these children are of marked superior ability. Although 60 per cent of the children are under age for their grades, their teachers are of the opinion that as a group they are not exceptionally bright and consider that in most instances their intelligence scores give them too high a mental rating.

The above results would seem to indicate that the opinion of Phillips,[1] although possibly an over-statement, has a large element of truth. He says, "I have demonstrated and can demonstrate that, of all the intelligence tests yet published, 75 per cent of the questions depend more on experience, on associations and on the general and specific education of the individual than on native intelligence. Of the 41 children found in California having a superior IQ all save one belonged to families of culture and intelligence, and of which one or both parents were college graduates. Association in environments, not native intelligence, qualified them to answer."

The writer (C.) is moved to add in comment the following: *All* intelligence tests yet published depend for their validity on experience, on association and on the general and specific education of the individual. However (and here must be added the qualifying statement to prevent those who read these opinions from making a wrong inference as to the worth of these tests), such tests, depending as they do *absolutely* on the experiences of those tested, are valid in showing differences in native mentality when, *and only when*, those tested have had common experiences and similar interests.

From the evidence now in our possession we can reasonably conclude that linguistic ability is an important factor in the score obtained by an individual in an intelligence test that is based largely on words and their uses. This ability must definitely be considered whenever an individual in a group tested deviates in any marked degree in this ability from his fellows in the group. How to discover this deviation in linguistic ability is, however, in most instances a difficult task. Of course when the individual examined is foreign born, and when he speaks English in a halting and inaccurate way, he can be easily singled

[1] *American Education*, Vol. XXVI, No. 1922, p. 61.

out for special attention. It has been pointed out, however, that an individual may possess a pronounced linguistic handicap and still speak the vernacular with a reasonable accuracy and fluency. It is also to be remembered that the possession of unusual linguistic fluency may give an individual a decided advantage in an intelligence test and result in the securing by him of a score that places his apparent mentality considerably above what it really is. Such cases as these clearly cannot be detected by ordinary observation. Various ways suggest themselves as means by which a linguistic handicap or unusual verbal fluency may be discovered.

1. In testing a group of children or adults to determine their mentality, a preliminary test to discover their reading ability might be employed. For instance the various tests devised by Thorndike to measure the extent of vocabulary and the understanding of paragraphs read silently, suggest themselves. However, such tests would help us in our own particular problem little, if at all, and for this quite apparent reason—the differences in reading ability discovered might be due to one of two causes, or to both combined. The differences in the scores obtained by various individuals might result from actual differences in intelligence, or from differences in opportunities to acquire familiarity with and fluency in the vernacular. Without further evidence there would be no way of determining the presence of these two factors or their relative amounts.

2. A preliminary survey might be made of the social and economic status of the individuals tested. When this status is decidedly below that of the average of the group, there would be presumptive evidence of a significant language handicap. In particular, children and adults living with foreign born and largely non-English speaking groups should be singled out for further investigation. On the other hand those coming from particularly favored environment where the social status is high and the cultural influences superior should also be given special consideration, since their intelligence scores are likely to indicate a mentality in excess of that which they actually possess.

3. In testing a group it would be well to employ a number of mental examinations. Our belief is that for children in the upper grades, youth in the high school, and adults in our colleges and universities, the procedure should be somewhat as follows: There should be two tests of the linguistic type given on succeeding days, or weeks under conditions as to time, place, and methods of administration as nearly identical as possible. These tests may be either two forms of the same

mental examination, or two different examinations, preferably the latter. In any event each of these examinations should be preceded by a simple fore-examination to acquaint all with the nature of the tests to be employed. Later, a third mental examination of the performance type should be given. When there is substantial agreement for a given individual in the results of these separate examinations then the scores may be taken as expressing reasonably well the actual mentality of that individual. When there is a marked disagreement this means that such an individual will require further investigation, such as interviews and testing by individual examinations of the Binet and performance types. In this connection it should be pointed out that we need better developed and more completely standardized tests of the group performance type, adapted to the intelligence of older children and adults. The majority of the performance tests so far devised have been prepared for children in the primary grades, illiterates and non-English speaking adults. These are not difficult enough, as a rule, to discriminate between the higher levels of mentality. One test, however, the Myers' Mental Measure,[1] is a group examination entirely of the non-linguistic variety, designed for all grades from the kindergarten to the university. Dearborn has likewise extended his Non-linguistic Tests to cover all grades through the high school. The writer (C.) has not employed this test in his own work, but published accounts seem to indicate that it has been reasonably successful in measuring native intelligence.[2]

4. When intelligence tests are employed primarily for the purpose of promotion and elimination, particularly for the latter purpose, no pupil or student who scores conspicuously low should be dealt with entirely on the basis of such mental examinations. These pupils and students should be further investigated. Their previous academic records should be studied, the opinions of their teachers obtained,

---

[1] *Twenty-first Yearbook* of National Society for the Study of Education, Chapter IV. Public School Publishing Co.

[2] Perhaps a note here may not be out of place in regard to the relation between non-verbal tests and linguistic ability. It does not seem to follow necessarily when a test is employed in which no words are used that for this reason linguistic ability is not brought into play, particularly when rational processes are involved. The writer (C.) constantly catches himself when working with non-verbal tests using inner speech which becomes more and more conscious when a problem increases in difficulty and complexity. This seems to mean that we cannot think to any extent without the use of words, hence we can never hope to frame intelligence tests that do not depend to a certain degree upon linguistic knowledge and ability.

their social and economic status investigated, etc. There should be at least one personal interview, and a further mental examination of the individual type given.

From the results that have been obtained, particularly with students at Brown University, based on investigations now extending over a period of more than 4 years, the writer (C.) is convinced that the group examination of the verbal type when given with care will reveal in general the actual mentality of those tested to a degree of accuracy that is sufficient for all practical purposes in from 80 to 90 per cent of the cases tested.

That it fails at times to indicate accurately the real mentality of the individual or a group of individuals does not mean that it should be abandoned, but it does mean that in cases vitally affecting the school career of the pupil utmost caution should be employed.

TABLE I.—SUMMARY

	AMERICAN	ITALIAN
1. Number of pupils	50	50
2. Grades	5–8	5–8
3. Ages	11–16	11–15
4. Average chronological age	13.12	13.11
5. Average pedagogical age	11.77	12.
6. Average score (National)	103	90
7. Average IQ.		
Terman	92	91
National	85	76
8. Average score-National Test.		
Part 1	25	24
Part 2	20	17
Part 3	21	17
Part 4	10	10
Part 5	27	22
	103 (American	90 (Italian)
9. Correlation between two tests ρ =	.73 (American)	.79 (Italian)

TABLE II.—AMERICANS

Pupil	N	T	D	Pupil	N	T	D
1	15.5	15	− 0.5	26	11	12	1
2	15	15	0	27	11	12	1
3	14	12	− 2	28	10	13	3
4	14	12	− 2	29	10	12	2
5	13.5	13.5	0	30	10	10	0
6	13.5	14	.5	31	10	11	1
7	13.5	14	.5	32	10	13	3
8	13	12	− 1	33	10	11	1
9	13	12	− 1	34	10	13	3
10	13	14	1	35	10	12	2
11	13	12	− 1	36	10	11	1
12	12.5	12.5	0	37	10	10	0
13	12	12	0	38	10	10	0
14	12	13	1	39	10	12	2
15	12	12	0	40	10	12	2
16	12	13	1	41	10	11	1
17	12	12	0	42	9	10	1
18	12	13	1	43	9	10	1
19	11	13	2	44	9	11	2
20	11	11	0	45	9	10	1
21	11	11	0	46	9	10	1
22	11	11	0	47	8.5	10	1.5
23	11	12	1	48	8.5	11	2.5
24	11	12	1	49	8	11	3
25	11	12	1	50	8	10	2
			10				+38
			−7.5				
			+2.5				

TABLE III.—ITALIANS

Pupils	N	T	D	Pupils	N	T	D
1	15.5	14	− 1.5	26	9.5	12	2.5
2	15	15	0	27	9.5	13	3.5
3	14	13	− 1	28	9.5	11	1.5
4	14	13	− 1	29	9	11	2
5	12	12	0	30	9	12	3
6	12	14	2	31	9	14	5
7	11	13	2	32	9	11	2
8	11	11	0	33	9	11.5	2.5
9	11	12	1	34	9	11	2
10	11	12	1	35	9	12	3
11	11	12	1	36	9	11.5	2.5
12	11	11	0	37	9	11	2
13	11	11	0	38	9	12	3
14	11	11	0	39	9	11	2
15	11	12	1	40	9	12	3
16	10.5	11	.5	41	8.5	10	1.5
17	10.5	11	.5	42	8.5	11	2.5
18	10.5	11	.5	43	8	11	3
19	10.5	11	.5	44	8	11	3
20	10	11	1	45	8	10	2
21	10	12	2	46	8	10	2
22	10	11.5	1.5	47	8	10	2
23	10	11	1	48	7.5	10	2.5
24	10	11	1	49	7.5	10	2.5
25	10	11	1	50	7.5	10	2.5
			17.5				+63
			− 3.5				
			+14.				

TABLE IV.—TABLE SHOWING THE RELATION BETWEEN SCHOOL GRADE AND NATIONAL SCORE

50 Italians

Scores	Grade and pedagogical age						Totals
	V*A* 10.5	VI*B* 11	VI*A* 11.5	VII*B* 12	VII*A* 12.5	VIII*B* 13	
130–143	..	..	..	1	..	..	1
120–129.9	..	..	..	..	1	2	3
110–119.9	..	..	..	..	..	2	2
100–109.9	..	..	..	2	5	2	9
90– 99.9	..	..	3	4	3	..	10
80– 89.9	..	..	1	4	4	..	9
70– 79.9	..	2	3	3	1	..	9
60– 69.9	1	5	1	..	..	..	7
Totals.....	1	7	8	14	14	6	50

50 Americans

Scores	Grade and pedagogical age						Totals
	V*A* 10.5	VI*B* 11	VI*A* 11.5	VII*B* 12	VII*A* 12.5	VIII*B* 13	
130–144	..	..	..	1	2	..	3
120–129.9	..	1	1	1	2	3	8
110–119.9	..	..	1	..	1	5	7
100–109.9	..	1	1	2	1	..	5
90– 99.9	..	7	4	3	2	..	16
80– 89.9	1	3	1	..	..	..	5
70– 79.9	..	4	..	..	..	..	4
60– 69.9	..	2	..	..	..	..	2
Totals.....	1	18	8	7	8	8	50

TABLE V.—COMPARISON OF ACCOMPLISHMENT OF 173 AMERICANS AND 163 ITALIANS IN THE VA GRADE, AS MEASURED BY THE LIPPINCOTT-CHAPMAN CLASSROOM PRODUCTS SURVEY TESTS

Arithmetic Fundamental Test

Score	Distribution of scores of 173 Americans (Each star indicates an individual score)	Totals
Highest—9	**********	15
8		0
7	**********	21
6	**********	46
Median—5	**********	52
4	**********	24
3	*********	9
2	******	6
Lowest—1		0
	Grand total	173

Score	Distribution of scores of 163 Italians (Each star indicates an individual score)	Totals
Highest—9	**********	11
8		0
7	**********	22
6	**********	33
Median—5	**********	46
4	**********	29
3	**********	14
2	*****	5
Lowest—1	***	3
	Grand total	163

TABLE VI.—COMPARISON OF ACCOMPLISHMENT OF 173 AMERICANS AND 163 ITALIANS IN THE VA GRADE, AS MEASURED BY THE LIPPINCOTT-CHAPMAN CLASSROOM PRODUCTS SURVEY TESTS

Arithmetic Problems Test

Score	Distribution of scores of 173 Americans (Each star indicates an individual score)	Totals
Highest—9	**********	34
8	**********	23
7	**********	34
6	**********	34
Median—5	**********	36
4	**********	10
3		0
2	**	2
Lowest—1		0
	Grand total	173

Score	Distribution of scores of 163 Italians (Each star indicates an individual score)	Totals
Highest—9	********	8
8	**********	10
7	**********	22
6	**********	48
Median—5	**********	41
4	**********	22
3	********	8
2	***	3
Lowest —1	*	1
	Grand total	163

TABLE VII.—COMPARISON OF ACCOMPLISHMENT OF 173 AMERICANS AND 163 ITALIANS IN THE *VA* GRADE, AS MEASURED BY THE LIPPINCOTT-CHAPMAN CLASSROOM PRODUCTS SURVEY TESTS

Reading Selections Test

Score	Distribution of scores of 173 Americans (Each star indicates an individual score)	Totals
Highest—9	**********	31
8	**********	14
7	**********	24
6	**********	43
Median—5	**********	46
4	**********	10
3	***	3
2	*	1
Lowest —1	*	1
	Grand total	173

Score	Distribution of scores of 163 Italians (Each star indicates an individual score)	Totals
Highest—9	*	1
8	*	1
7	***	3
6	**********	10
Median—5	**********	33
4	**********	39
3	**********	30
2	**********	28
Lowest —1	**********	18
	Grand total	163

TABLE VIII.—COMPARISON OF ACCOMPLISHMENT OF 173 AMERICANS AND 163 ITALIANS IN THE VA GRADE, AS MEASURED BY THE LIPPINCOTT-CHAPMAN CLASSROOM PRODUCTS SURVEY TESTS

Reading Continuous Passage Test

Score	Distribution of scores of 173 Americans (Each star indicates an individual score)	Totals
Highest—9	**********	39
8	**********	27
7	**********	36
6	**********	24
Median—5	**********	24
4	****	4
3	********	8
2	****	4
Lowest —1	*******	7
	Grand total	173

Score	Distribution of scores of 163 Italians (Each star indicates an individual score)	Totals
Highest—9	**	2
8	***	3
7	**********	14
6	**********	14
Median—5	**********	48
4	**********	28
3	**********	18
2	**********	15
Lowest —1	**********	21
	Grand total	163

TABLE IX.—COMPARISON OF ACCOMPLISHMENT OF 173 AMERICANS AND 163 ITALIANS IN THE VA GRADE, AS MEASURED BY THE LIPPINCOTT-CHAPMAN CLASSROOM PRODUCTS SURVEY TESTS

Total Test Score

Score	Distribution of scores of 173 Americans (Each star indicates an individual score)	Totals
Highest—9	**********	51
8	**********	27
7	**********	31
6	**********	22
Median—5	**********	16
4	**********	16
3	*********	9
2	*	1
Lowest —1		0
	Grand total	173

Score	Distribution of scores of 163 Italians (Each star indicates an individual score)	Totals
Highest—9	****	4
8	***	3
7	********	8
6	**********	14
Median—5	**********	32
4	**********	37
3	**********	39
2	**********	19
Lowest —1	*******	7
	Grand total	163

THE EFFECT OF ENCOURAGEMENT AND OF DISCOURAGEMENT UPON PERFORMANCE

GEORGINA STICKLAND GATES

AND

LOUISE Q. RISSLAND

Barnard College, Columbia University

Instructions given to persons learning to administer psychological tests usually include the advice, to praise and encourage the subject whose ability is being investigated. Terman,[1] for example, says: "Nothing contributes more to a satisfactory *rapport* than praise of the child's efforts. Under no circumstances should the examiner permit himself to show displeasure at a response, however absurd it may be. Exclamations like fine, splendid, etc., should be used lavishly."

The general assumption that encouragement has a favorable and discouragement an unfavorable effect on performance is borne out by an experiment of Gilchrist's.[2] He found that one group of 29 subjects to whom he said, after giving a test, "A hasty examination of the papers in the test just given shows that the members of this group did *exceptionally well.* I ask you to take the test again;" improved 79 per cent in the second test, whereas a group of 21 subjects to whom he said, "A hasty examination of the papers in the test just given shows that the members of this group *did not do so well in the test as the average 12-year-old child would do.* I ask you to take the test again;" made a 6 per cent lower group score in the second examination than in the first.

In the present experiment an attempt has been made to investigate further the effect of the experimenter's comments on two very simple performances. The subjects used were 74 college students who were given individually, after a preliminary exercise, two trials of the Motor Coordination (Three Hole) and two of the Color-naming Test. After taking the first coordination test, the first subject was told, "That is really splendid! Do you always make such good scores? In a curve of distribution your score would be way up here (indicating a position at the top of the curve). Your score was so good that I wonder if you would mind repeating the test?" After taking the test

[1] Terman, L. M.: "The Measurement of Intelligence." 1916, p. 125.

[2] Gilchrist: The Extent to Which Praise and Reproof Affect a Pupil's Work, *School and Society,* Dec. 2, 1916.

again and after performing the first trial of color-naming, she was encouraged similarly with words and inflections which had been previously standardized. To the next individual who took the coordination test, the experimentor said, "Oh dear, that is really a very poor score. I am afraid that you would fall at the bottom of the curve of distribution," etc. Expressions of disappointment and sympathy were similarly offered at the completion of the first color-naming test. To one-third of the group no comment concerning their performance was made; they were simply asked to repeat the test.

Certain obvious precautions were observed. The subjects promised not to tell any other persons about the experiment. They were asked to write down what they believed the purpose of the investigation to be. Only two suspected the object; the others giving such replies as "To find the correlation between motor coordination and visual perceptions;" "To find the speed with which one can react to stimuli," etc. The experimentor had had considerable practice in college dramatics and that her performance was a convincing one was evidenced by the genuinely elated or disappointed demeanor of the subjects.

Although the selection was made purely on the basis of the order in which the subjects chanced to come to the experimentor, the three groups were found to be, as is shown by Table I, approximately equal in initial ability. Table II shows the average improvement of the individuals working under different incentives, and Table III the per centage of each group which improved, fell off, or remained the same. From Table II, it appears that the difference in average improvement is slight, and that the SD's are relatively large. The direction of the difference is not the same for both tests; in the Coordination Test the order from greatest to least improvement is, encouraged, discouraged, repetition; in the Color-naming Test the discouraged group made the greatest gain, the encouraged next, and the individuals who merely repeated the test became less proficient. The small percentage of improvement under encouragement (9 per cent in one test and 0.007 in the other) contrasts with the 79 per cent found by Gilchrist; similarly the improvement found under discouragement (6 per cent and 0.01) is in disagreement with the deterioration of 6 per cent found by the former writer. The percentage of individuals who improved (see Table III) is, in both tests higher for the encouraged than for the discouraged group, and higher for the latter than for the repetition group. But the differences are again small; a change of a

few individuals from improvement to deterioration or *vice versa*, would make a difference in the order of the hierarchy.

Though the groups in Table IV are necessarily small, an attempt is made there to show whether the various incentives have any differential effect upon groups unlike in initial ability. The variations in the table appear to be chance ones differing with the test, with the possible exception of the effect of discouragement upon the poorest group of subjects. Whereas under encouragement, in coordination, 100 per cent, and in color-naming 63 per cent of the poorest subjects improve, under discouragement only 57 per cent and 29 per cent make better scores in the second test. The average improvement of 10.8 taps of the encouraged group contrasts with the gain of 6.9 taps of the discouraged group in Coordination, and the gain of 5.1 seconds contrasts with the loss of −0.11 seconds in Color-naming. Though the apparent difference may well be a spurious one, it recalls another quotation from Terman.[1] "In general, the poorer the response, the better satisfied one should appear to be with it." Possibly this apparent effect of discouragement partly explains some of the parts of Table V where the negative correlations so frequently obtained between initial ability and improvement appear, except in the case of discouragement where the tendency is toward a zero or slightly positive relation.

It is evident from Table VI that the effect of encouragement or discouragement (if such an effect exists) is not necessarily the same for the same individual in both tests. In 50 per cent of the cases, individuals who improved in one test after being praised or sympathized with, did not improve in the other. The same phenomenon of a larger number of individuals improving under encouragement and discouragement than under repetition, appear in this table as it did in Table II and III.

The results of this study seem to show, then, a very slight difference in average improvement or even in percentage of individuals who improve in the three groups. In this, the facts found are similar to those observed in experiments on fatigue, lack of fresh air, sleep, or food; the external factors seem to be of relatively little importance in determining the score. Such difference as there is seems to be in favor of encouragement or discouragement rather than mere repetition. We might say then, (with the usual realization of the inadequacy of the data) that it is better to make some comment about the score than

[1] Terman: "The Measurement of Intelligence." 1916.

to make none; that it is a little better to make an encouraging than a discouraging remark; that relatively poor individuals are more likely to be unfavorably affected by discouragement than are relatively proficient persons; that the effect of these incentives does not seem to be constant for the two tests. The desirability of performing such an experiment on more susceptible subjects, as children, using more complex, and more reliably measured functions, is obvious.

TABLE I.—AVERAGE PERFORMANCE OF THREE GROUPS IN INITIAL TEST

Group	Coordination		Color-naming	
	Average number of taps per minute	AD	Average time in seconds	AD
Encouraged	94.6	10.6	53.5	6.4
Discouraged	92.5	10.4	57.0	7.9
Repetition	94.7	10.2	55.2	8.7

TABLE II.—AVERAGE IMPROVEMENT OF THREE GROUPS

	Coordination		Color-naming	
	Number of taps per minute	SD	Time in seconds	SD
Encouraged	8.7	7.1	0.4	4.3
Discouraged	5.4	9.9	0.7	4.5
Repetition	4.9	7.6	−1.4	5.1

TABLE III.—PERCENTAGE OF SUBJECTS WHO IMPROVED, FELL OFF, OR REMAINED THE SAME IN THE THREE GROUPS

	Improved	Fell off	Remained the same
Coordination			
Encouraged	89	11	0
Discouraged	70	26	.04
Repetition	64	28	.08
Color-naming			
Encouraged	58	38	.04
Discouraged	51	40	.09
Repetition	44	48	.08

TABLE IV.—THE EFFECT OF INCENTIVES UPON GROUPS DIFFERING IN INITIAL ABILITY

Coordination	Average improvement in taps	Percentage of subjects who improved	Percentage of subjects who fell off	Percentage of subjects who remained the same
Encouraged				
Highest 8 subjects....	5.9	63	38	0
Middle 10 subjects....	11.5	100	0	0
Lowest 8 subjects.....	10.8	100	0	0
Discouraged				
Highest 7 subjects....	2.0	72	28	0
Middle 9 subjects.....	8.4	89	0	11
Lowest 7 subjects.....	6.9	57	43	0
Repetition				
Highest 8 subjects....	3.6	63	37	0
Middle 9 subjects.....	2.0	45	33	22
Lowest 8 subjects.....	9.4	87	13	0
Color-naming				
Encouraged				
Highest 8 subjects....	.5	75	25	0
Middle 10 subjects....	1.0	50	50	0
Lowest 8 subjects.....	5.1	63	25	12
Discouraged				
Highest 7 subjects....	2.9	57	43	0
Middle 9 subjects.....	1.9	67	22	11
Lowest 7 subjects.....	− 1.1	29	57	14
Repetition				
Highest 8 subjects....	− 2.9	37	63	0
Middle 9 subjects.....	− 1.0	55	33	12
Lowest 8 subjects.....	2.5	38	50	12

TABLE V.—CORRELATION BETWEEN INITIAL ABILITY AND IMPROVEMENT

	Coordination	Color-naming
Encouraged	−33	−24
Discouraged	−.03	+.16
Repetition	−34	−20

TABLE VI.—RELATION BETWEEN IMPROVEMENT IN ONE FUNCTION AND IMPROVEMENT IN THE OTHER

	Encouraged	Discouraged	Repetition
Percentage of group which improved in both tests	46	37	20
Percentage which became worse in both tests	4	13	16
Percentage which improved in one test and not in the other	50	50	64

WHAT SHALL WE EXPECT OF THE AQ?

HERBERT A. TOOPS

AND

P. W. SYMONDS

Institute of Educational Research

Teachers College

(*Continued from December*)

Other Causes of Poor School Work Besides "Motivation."—The AQ technique was not designed as a "cure-all." But humanity, generally, seems to prefer to rely upon one simple technique, such as the AQ procedure, than to attempt to analyze its ills in terms of more complex, and numerous causes. It can thus do no harm to point out at this juncture a number of other factors that are in the aggregate perhaps quite as important as "motivation." This may be readily shown by making evident the fact that at best the AQ procedure *merely singles out the pupil who thereupon needs diagnosis and treatment* therefor, or an application of more "motivation" not primarily designed to remedy a remediable defect. It does not diagnose the *cause* of poor school work.

Both Pintner and Franzen seem to recommend the AQ procedure as an administrative device to be undertaken particularly *during the course of the academic year* to discover either those pupils, those classes, those grades, or those schools which are not working hard enough. For those pupils "who are not working hard enough" incentive to improvement may be provided, and a very powerful one, as above shown, especially when a graphic method is employed for showing to a child his efforts as measured by himself. If a class is found which is not doing work "up to par," evidently the conclusion would be that the teacher is not doing her duty, or that home or other conditions are to blame. If a grade is not up to par, the conclusion might be that the textbook is at fault, the curriculm to blame, or the general educational methods of the grade at fault. And if an entire school is below par, the causes might be harder to discover. Pintner cites other causes of classes doing poor work such as: Crowded conditions, poor attendance, inadequate lighting, undernourishment, or bad physical conditions of pupils. A current newspaper report of a Chicago investigation of the bearing upon scholarship of movie attendance discloses the fact that of 3000 pupils in six

schools, the 275 "best" pupils used in one week an average of 1.43 movie tickets per pupil, while in the same week the 275 "poorest" pupils used an average of 1.83 tickets per pupil. Again, Reavis finds high negative correlations between school achievement and distance traveled to school by rural pupils.[1]

An Analysis of the Elements Involved in School Placement.—If we believe that results thus far obtained from the AQ procedure are due in large part to the ultimately less than unity correlation of mental age and educational age, the educational problem then resolves itself into those of: (1) Proper placement or sectioning at the outset; (2) special educational remedial measures to provide the "tools of learning" (reading ability in many cases, special arithmetical processes in others) for those who are doubly promoted or are over-promoted with respect to their section (provided such pupils are not to be demoted) for the present; (3) measurement of educational product at frequent intervals during the academic year, (4) remedial measures and incentives for those not making the "standard" amount of progress for pupils of that section, due to other causes than over-promotion. The third step, the measurement of educational product, will ultimately take the place of the AQ procedure. This follows from the fact that with adequate sectioning, the denominator of equation (1) is practically the same for all people in a given grade, since proper sectioning means exactly that thing: *Delimiting the range of capacity for progress* so that all pupils of a given class have approximately the same capacity for progress. Thus ideal sectioning, once set going, automatically becomes the AQ procedure; but with the hypothesis the converse statement is not true, since the present IQ on any test cannot hope to be the best possible measure of capacity for progress: The AQ procedure once set going does not become ideal sectioning. This suggests that in the long run, the AQ may be unable to prove in practice as good a motivator as we might hope. The present correlation of EQ and IQ is certainly less than 0.87 in all school subjects; with that high a correlation, the standard error of estimate in predicting EQ from MA is half as large as the standard deviation of the MA's. And finally we must realize that all, the dull and the bright, by easily created artificial motives of which the AQ is one of the most potent and satisfactory yet brought to the attention of the educational world, can be brought up to a point of educational attainment far

[1] Reavis, G. H.: Factors Controlling Attendance in Rural Schools. *Teachers College Contributions to Education*, No. 108, New York, 1920, 68 pp.

beyond what we have thus far thought possible. Writers, philosophers, have defined education as a continual readjustment process. It is a commonplace to remark upon the fact that a boy placed in school (sectioned) on the basis of IQ "does not stay put." Our promotions must come oftener so that the remedial measures will be less necessary. It will probably always be necessary for promotions to go by whole steps, by jerks and jumps, rather than at a steady rate. Unless educational offerings become steadily more difficult in exact parallel with developing mentality and developing willingness to work,[1] then educational promotion must be jerky and remedial measures will be required to provide the tools of learning for the promotee. In the army E. and R. schools, promotions in certain courses came every 2 weeks provided the pupil passed a certain score on a standard test. If the pupil failed to make the promotion he was retarded but 2 weeks. Many soldiers made three or four school grades of progress in reading in 6 months by this scheme. This leads us to believe that probably the school problem where the AQ procedure is applicable is largely one of initial placement, or sectioning. It is also suggested that a mentally bright pupil now in Grade VI but suddenly promoted to the Grade VIII may not at once begin to do there the work required of him. It requires a distinctly higher level of *reading ability* to do the work of the Grade VIII than Grade VI. Not subjected in Grade VI to the Grade VIII reading requirement the doubly promoted pupil is lacking one of the essential elements or tools for progress in the suddenly confronted new situation of the new grade. From this we may conclude that ability to get along in school work depends upon *at least* four generalized things: (1) Willingness to work, effort, motive, or interest; (2) mental ability, as may be measured by some sort of general intelligence test; (3) previous preparation, one of the most important components of which is reading ability as shown by Osborne,

[1] At the present time the grades become successively more difficult relative to mental capacity. This is shown by by the facts of school elimination and the higher average IQ's in successively more advanced school grades. There are practically as many individual rates of growth of intelligence as there are numbers of individuals. A school grade must be adapted in difficulty to some given M.A. level. The people of this M.A. level will have different rates of growth of intelligence, so that when all may have the same M.A. at the beginning of a semester, there will be large variations at the end of the semester. We may hope that our schools of the future will come more and more to provide individual rates of school progress for each pupil, such as is often now attempted in certain business colleges and trade schools.

Mrs. Burgess and others; and (4) physical capacity. If pupils were sectioned at the beginning of every semester on the basis of a properly weighted composite scale made up of the four elements above, the present correlation between "capacity" and educational age would be very greatly increased, and with this change in r would disappear most of the cases of "overworked or underworked" pupils. If finally, the schools were immediately to promote in school 1 year, 2 years, etc., just as large a percentage of the mentally bright as of the mentally dull who are now retarded in school, and give these advanced pupils special drill in reading and the drill of previous preparation required to do today's problems, in a few months the correlation would be still further raised toward unity, the goal of the AQ workers.

How can a bright pupil who by mental ability "belongs" in the school grade which is two grades in advance of the school grade which is normal for one of his chronological age be expected to obtain the educational progress of the 2-years-advanced grade if he is not allowed to go into that grade? Concretely, Johnny, chronological age of 10, is kept because of his age in Grade IV which is normal for a child of 10; he has a mental age of 12, which means that he is a bright child, of mental capacity such that he should be able to do work of about the Grade VI. Now the AQ hypothesis assumes that he should have the educational age 12 because he has the mental age 12. But how can Johnny get an educational age of 12 if he has not been subjected, and for some time, to the kind of educational content (Grade VI content) to which mentally normal 12-year-olds are normally subjected? It is of course probable that even with ideal sectioning and the best remedial treatment, a few pupils would refuse to work up to capacity. Such cases would, however, be pathological cases, rather than cases requiring normal treatment; for, as has been shown in ideal sectioning the AQ procedure becomes merely measurement of product for all normal (*i.e.*, non-pathological) cases.

Some Elements in Ideal Sectioning.—If better sectioning, involving increased rates of promotion, is so desirable, how shall we secure it? By the use of intelligence tests, has been the answer of the past. But these are inadequate, although far better than the system of personal judgment which they are fast replacing. Intelligence tests, and particularly group tests, in the past have meant an expression almost synonymous with "ability to make the average school marks as they are now made in our public elementary schools." They predict these

but poorly. This may be a blessing in disguise as will be pointed out below in our discussion of the need for a socialized curriculum.

Accurate placement sectioning as will be shown by one of the authors in a forthcoming article, is best to be secured not by a general intelligence test, or by a general intelligence test plus an educational progress test, but rather by a scale involving both factors and more, all factors being weighted so that the composite makes up a scale for best predicting a *specific subject ability*. In the long run, it means that best educational results would be obtained by sectioning with regard to each separate educational process as measured by a scale of specific application to that educational process. It points to a possible discard of the "general" scales and the adoption of specific scales. Since general ability is a concept of "averageness," or ability to make average school grades, and since the numerator of equation (1) is most general when we include in it the average subject-matter of many subjects, we should expect, on statistical grounds alone, that the more general accomplishment quotient will yield fewer extreme ratios than the more specific subject-matter accomplishment quotients. But subjects are not taught *en generale;* everywhere the pupil recites, now in reading, now in arithmetic—that is, in a specific subject-matter. The teacher of arithmetic is at present concerned mainly, if not solely, in how well or how poorly the given pupil can do arithmetical problems, and in the methods required in *his* case to teach *him* to work such problems faster and more accurately. Of what avail for her purposes to know that Johnny is up to par in "general effort" if "he simply isn't working" in arithmetic? In matters of more general administrative policy, the more general accomplishment quotients may be of more value. Teaching involves the problem of adaptation of method to the *specific* individual; school administration deals more with matters of general policy. In many individual cases, the general AQ as a measure of general school attitude, may be of use to the teacher, for it may be the key to the discovery of the lack of motivation in a specific school subject, as, for instance, in the case of a boy of good ability who has acquired a general distaste for all school work. Motivation in special subjects undoubtedly correlates highly with general school attitude, although higher in some school subjects than in others. Good conduct and good marks in the grade school generally carry over into the high school. Where lack of motivation is due to bad home influences, the general AQ will be of more value than in the case of those pupils whose maladjustment is subjective or

due to specific school conditions. In the future, we may expect much more attention to be devoted to methods of learning and methods of teaching specific subject-matters than to the problems of learning and teaching in general.

The problem of what tests to use in sectioning thus resolves itself in the last analysis to a question of what tests *will best predict the rates of ability to progress in acquiring those things which we wish to be acquired.* This suggests that instead of immediately looking for an ultimate measure or rate of getting along in school work, we should instead look first for ultimate standards of desirability of what one should obtain from education in order that we may recognize, by the correlations, a good measure of "intelligence" when we find it. Researches into what we should obtain from our school studies are sorely needed. If we had an ideal course of study, then "ability to make school marks" might be made quite synonymous with "intelligence," or with "school intelligence" at least. We must not conclude, however, that the only worthwhile adjustment of the curriculum is the adjustment of its difficulty to intelligence as our tests now define it. The curriculum must be worked out on the basis of social needs, whatever relation it may have with the verbal, abstract intelligence needed to cope with it. Some pupils may even be totally unable to profit by the one elementary school curriculum; for them a differentiation of subjects is desirable in order that such talent as they have may function to the best social advantage. Still others may profit maximally from less than the total amount of the present curriculum.[1] We shall soon be compelled to answer the question of whether or not it is desirable for a graduate of the Grade VIII to enter high school when the chances of his being graduated are very low; that is, when the likelihood of failure during the first few years of high school is very great. In a society such as ours, the "lowest 2 per cent" or more, in any "general" ability may consider themselves badly misplaced, "undesirable," fit subjects for replacement in our social-vocational scheme by improved automatic machinery, or by removal to a less desirable social class, by vocational elimination, or by segregation. These are the facts, unjust to those unfortunate individuals as, superficially, it may seem to be. While the sentimentalist is trying

[1] See the discussion of "upper auxiliary classes," Davis, H.: The Use of Intelligence Tests in the Classification of Pupils in the Public Schools of Jackson, Michigan. *Twenty-first Yearbook* of the National Society for the Study of Education, Bloomington, Ill., 1922, pp. 134–135.

meanwhile to decide the "justice" of the case, the school administrator must daily be doing something about it.

Rapid Promotion[1] *as a Possible Substitute for the AQ Procedure.*—Let us also consider the bearing of rapid promotion upon the problems of motivation, retardation and elimination. Would not a division of the school course into 50 grades instead of 8 provide, by the rapid promotions thus possible, the motive that is now seen to be efficient in the results derived from posting or plotting the AQ's. It is so worked out in the army schools above reported, and in a limited number of school subjects, it is true. But is it too much to expect that other school subjects will work out in the same way? With sectioning based on ability to make progress in such courses of the "ideal social values" in subjects, we shall approach perfect correlation between "intelligence" and school ability. Pupils able to make the hurdles frequently will be induced to stay in school with fewer compulsory education laws. Wooley has said, and has been since substantiated by others, that: "Failure, not economic necessity, is the cause of two-thirds of our elimination from school before the eighth grade is reached." This plan means *promotion by subject.* Why fail the pupil for a semester because he fails in arithmetic? To such a question one hears the plaintive reply: "But if all the above were desirable which isn't yet proven, how could it be administered?" The answer is that 50 grades does not necessarily mean 50 rooms nor 50 teachers, but only 50 times for promotion in the elementary school course. The departmental plan of instruction is being tried out in a number of New York City schools. Rapid promotion carries with it its own incentives.[2] The administration of this plan can best be carried out in a very large school. If such could be administered—and if proven desirable, then it becomes the problem of the administrator, or educational engineer, to devise the means—it would secure maximal effort of pupils. A pupil in such a system enters a given grade only because his intelli-

[1] Some educators are in doubt concerning the advisability of rapid promotion, advocating rather an expansion in topics covered for the pupils of superior ability. Franzen points out in his forthcoming dissertation that this is essentially a makeshift device. For a published discussion of minimum and maximum essentials as related to the AQ see Murdock, K.: The Accomplishment Quotient: Finding and Using It. *Teachers' College Record*, Vol. 23, No. 3, p. 22, pp. 229–239.

[2] Someone is said to have chided Napoleon for giving so many "toys" to his soldiers in the shape of medals. "Toys are these?" replied this thorough understander of human nature, "Well, men die for toys." It is possible to maintain the individuality of modern teaching and yet provide the "toys" of frequent promotion.

gence, previous training and physical capacity are such as to enable him to do successfully the work with pupils *of his own kind;* failure to gain promotion would be *self-admittedly deserved social disgrace.* But what is not possible in a modern metropolitan school where the number of students is thousands instead of hundreds?[1] And finally, is it too much to expect that our educational administrators will work out plans for *elimination from school by subject?* Cannot Johnny some day, under an enlightened educational system, hope to drop arithmetic only, instead of dropping out of school entirely?

We have discussed the value of the AQ as a "motivator," there are other methods of motivation fully as potent.

Some New Aims of Education and New Emphases in Methods Suggested by the AQ Procedure.[2]—The need in education for a measure of accomplishment, achievement, effort, motivation, attitude, dynamic effect of educational environment—call it what you will—is so great that it seems quite possible that the AQ procedure may be accepted uncritically by educators. Especially is this true in view of the fact that when the above methods are applied to large groups of pupils in a school system a very large majority of dull pupils are shown by the indices to be working hard, while a very large majority of bright pupils are shown by the indices not to be working hard. "*The bright pupil is the lazy one*" might easily become an educational shibboleth for quickly, and without proper consideration of what is likely to be the outcome, changing the point of emphasis of the teacher from the present one of paying particular attention to the dull pupils to that of paying particular attention to the bright pupils. The adoption of such a scheme in the school system would evidently give the teacher the necessary "scientific foundation" for relaxing her efforts with those unfortunate retarded pupils who popularly are already working "beyond capacity and will never come up to par in school work" and would allow her "conscientiously" to recommend them for vocational courses. This is highly desirable—granted, as it is very unsafe to do, that it can be assumed at once that such

[1] In New York City in 1921–22 there were 364 elementary schools with a population of 1000 or over; 164 with 2000 or over; 39 with 3000 or over; 5 with 4000 or over, and 1 with more than 5000 pupils.

[2] The thorough student of mental testing should read Bagley, W. C.: Educational Determinism: Or Democracy and the IQ, *School and Society,* Vol. 15, April 8, 1922, p. 373. This should be followed by the reply, Terman, L. M.: The Psychological Determist: Or Democracy and the IQ. *Journal of Educational Research,* Vol. 6, No. 1, 1922, pp. 57–62.

dull pupils cannot, with proper "effort," achieve a socially acceptable minimum of general education but can achieve a socially acceptable minimum of mechanical or vocational education. Again, the general adoption of such a measure might be expected to cause the teacher to redouble her efforts with the "lazy bright ones" who "are so much nicer to work with anyway," the geniuses of the next generation. This likewise is desirable—granted, as is likewise unsafe, that high AQ's in academic work are a method par excellence for developing the latent capacities of "genius."

If the AQ procedure results in mentally retarded students being advised in greater numbers than formerly to enter vocational courses, some good results will be obtained. Stenquist has shown that the correlation between mechanical ability and intelligence is not very high. This permits of a goodly number who are *below the average* in academic school work yet being *above the average* in mechanical ability. It emphasizes the adjustment phase of education; that education is not only individual adjustment, but readjustment, and makes more apparent the place of vocational education in a general educational system. Yet it must be noted that the correlations between intelligence and mechanical ability, while low, are always positive. This means that there are a considerable number of people in the world who are neither good at general education nor at mechanical work. Education cannot escape its duties in providing for these unfortunates by the AQ route, a process of "hand them on" if they work to capacity in academic education and yet get nowhere. A good axiom for school administration is to keep the child moving as long as he is growing mentally. It is absurd to keep a child in a grade 2 years, for "repeating work deadens interest and spoils ambition." As long as the child is developing mentally there is something he can learn; the solution lies in the line of differential rates of progress and differential curricula rather than in stricter compulsory education laws. Where shall the vocational school hand on the people who are neither good in academic work nor in mechanical work?

There is a still more serious count against the AQ procedure. The AQ implying perfect correlation of EQ and IQ, when *applied to specific* school subjects, would seemingly perpetuate the academic standard upon the most specific of subjects to which this procedure may be thoughtfully or thoughtlessly applied. Without caution we may be led to disregard the vast body of empirical evidence which points to a specialization of many abilities—lack of perfect correlation

among themselves—in different school subjects. If we are to use the AQ we must soon decide whether people studying art, mechanics, and automobile repairing, shall be required to come up to 1.00 AQ in those subjects; when obviously there may be little correlation between the EA's in those subjects and MA's determined by our present intelligence tests.[1]

G. I. Gates, following the suggestion of Hollingworth, has confirmed the belief that the correlation between functions increases as we approach the limit of practice in each. This lends strength to the hope that the correlation of EQ and IQ in the ordinary grade school subjects may be nearer unity than it now is. If pupils were ideally sectioned and subjected to an ideal course of study, promoted just as fast as ability to progress would warrant (absolutely individualistic promotion) then we could provide some different sort of graphic motivation process which would keep pupils working up to capacity. If we can increase the correlation of EQ and IQ towards unity, extreme AQ's will automatically tend to disappear. If pressure is continually applied to schools whose AQ is below 1.00, in the long run higher standards, new norms, will be set for our school work; *there will thus always* be 50 per cent who will need special remedial measures to bring them up to the ascending average, or norm. This would tend to perpetuate and increase our present academic trend rather than aid in discovering the social objectives of the curriculum. The present trend is good with reference to the past; it is poor with respect to what the future should demand of education. It is evident that ultimately we have need for absolute standards of school proficiency rather than the annually changing relative standards of position within a test score distribution, or arbitrary "linear" scale based upon age or grade averages, which we are now employing. We need tests of arithmetic of the objective nature of "What is the average length of time required by Johnny to multiply two-place numbers by two-place numbers with an accuracy of 99.9 per cent?" With such tests on hand, we shall be ready to abandon all annually varying relative measures such as sigma, percentile or T-score position, percentage or letter grades. With such absolute standards, the "passing mark" will be x per cent of what is "essential," whether that be attained in 2 months, 10

[1] Franzen states that the AQ procedure is valid in "each of the elementary school subjects," by which he probably implies only the conventional reading, arithmetic, history, geography, etc. He tested out only arithmetic, reading and language.

months or 2 years. An important advantage for the improvement of teaching will be secured, for then if two teachers have pupils of the same average ability for progress, the product will be a very objective measure of the teacher's ability to induce motivation and to secure results.

If the AQ method brings the increased flexibility of the promotions above shown to be so desirable, if it results in newer and increased emphasis upon more adequate sectioning at the beginning of each semester, if it makes it more evident that a child newly promoted may be deficient in the tools of education at that point, if it provides the real incentive which it promises in making lazy pupils work, then it is a very real and worthy object of our admiration. Herring and Gates have found that individual AQ's in a specific school subject can readily be raised to 1.35, or multiples of that figure, if based upon a general norm of all school systems. The truth is that probably everyone can work harder than he now does. The AQ in its philosophy does not connect itself with the problem of whether or not a child is doing more than he is *physically* to be expected to do. This is a problem of which we know only too little. In addition to giving many tests, let us determine some of the objectives of an educational system based on social needs, and then devise more specific tests which shall insure the attainment of those objectives. Some will object by saying that an IQ has certain merits in its very *generalness* as *vs.* the *specificness* of specific scales for predicting specific subject matter abilities. The IQ may correlate fairly well with capacity to progress in reading, arithmetic, geometry, with ability as officer in the army, as executive in industry, and *many* others. A specific weighting of those tests would predict any one of these abilities better than does the general weighting, but once so weighted would apply well to only the one, or at best two or three of the abilities. It is true that we may be forced in certain cases to retain the IQ procedure for reasons of economy, even though specific weightings would give better predictions of more specific abilities. It should be noted that, as soon as we decide upon a "criterion of intelligence" we cease attempting to measure general intelligence; we are then attempting to measure a specific ability—*that ability which the criterion measures.* Thus the determination of the criterion—what we are trying to measure, expressed in concrete objective units—logically comes first, and the construction of tests second. Without historically having had the tests first as we have had in the past we would not realize half as keenly as now the need of a

criterion. Thus our work as scientists logically leads us to become philosophers. Is it too much to hope that science will prove our best tool in our new role of philosopher?

An EA measured over a socialized curriculum, rather than an "effort" measure, is the measure of ultimate worth of school attainment. We should not lose sight of the fact that an IQ of 0.70, working at an efficiency, AQ = 1.00, is after all acquiring educational assets which will be valued by the world as only EQ's of 0.70 are valued. When such a boy with EQ of 0.70 becomes an adult, society primarily values him for his physical brawn and industry, and not for his "intelligence" or "education."

Thus our consideration of the AQ has solved nothing, although it has raised many interesting problems.

REVISING THE DEARBORN INTELLIGENCE EXAMINATIONS

WALTER F. DEARBORN

AND

EDWARD A. LINCOLN

Psycho-Educational Clinic, Harvard University

One of the advantages, if not the chief advantage, of the group as compared with the individual test of intelligence is the saving in time where large numbers are to be examined. A class of 40 or a group of several hundred may be examined in the time which it might ordinarily take to examine any one of the group individually. To secure this result with as little loss as possible in the reliability of the individual finding is, of course, one of the desiderata of group testing. Experiments have indicated that, other things being equal, the longer and more varied a test, the greater its reliability. In view of the time saving in the aggregate as compared with the individual tests, it would seem fair to allow a generous allotment of time to the group tester. His tests may be given in several sections and in short periods, and the added returns will furnish the experimental basis for the selection of the most effective tests for a more abbreviated series, if that proves desirable. The Dearborn Group Intelligence Tests were planned by the author with such considerations in mind. It was expected that they would be found long, especially in the hand of untrained and inexpert examiners. In the experience of the writers, when the duration of the periods of testing has been properly controlled they have not proved irksome to the examiners, much less to the examinees. Statistical studies have, however, shown that an abbreviation of the tests is possible without changing materially the results at present obtained. The present revision has, therefore, eliminated certain of the tests. A subsequent revision might well proceed on the basis of replacing these tests by others with the purpose of further extending the usefulness of the scale.

The basis for these eliminations and for various minor changes is described below.

The examiners and scorers had noticed that a number of the tests in the First Series were so easy that they seemed of little value in

differentiating pupils of various intelligence levels. It was determined to investigate this point, and so the scores of each separate test in Series I were distributed by half-year intervals, and the medians were found.[1] The results of this study appear in Table I, and they show that some of the tests do not contribute especially to the diagnostic value of the examination above the early years. This fact is even more strikingly apparent when expressed graphically as in Diagram 1. Here it is seen that the maximum score on the Color-form Test is reached at an early age, and that there is no significant differentiation after the seventh year. On the other hand, there are continued increments of the Page 2 score, and the maximum is not reached by any considerable portion of the pupils of the twelfth year. Since these latter tests and others in the Series appear to furnish sufficient differentiation in the early years as well as in the later years, it was decided to drop the Color-form Tests. The writers would suggest, however, that a series of tests beginning with such a one as the Color-form Test, or a simpler test, each of which reached plateaus at successively higher levels might furnish a better scale than a combination of tests which have a larger range of age differentiation, such as is now found, for example, in the Opposite Completion Test of the Second Series.

It would be possible, of course, to use the Color-form Test by increasing the number of items or by decreasing the time. The former plan, however, would lengthen the test, and the time limit would have to be reduced to something less than half a minute to get a proper differentiation on the test. This would make for inaccuracy.

Other considerations were taken into account in the elimination of some of the tests. The Map Test, for instance, is a very good test for the differentiation of the older pupils, but it was omitted from the revision because teachers found the scoring rather difficult.

After the revision of the examination, the scores on the revised form were found, and were correlated with the original scores. In one town, where 532 children in the first three grades had been tested, the Pearson Coefficient of Correlation was found to be 0.97. In another community where all grades through the eighth were tested, a group of 822 children under 13 years of age gave a correlation of

---

[1] In selecting the separate tests, pages 1 and 2 of GI were each taken as a whole. No smaller division seemed worth while, as each of these pages is made up of a number of short problems, no one of which receives a considerable score except the Ball and Field Test.

0.99 on the two forms.[1] These high correlations indicate that the short examination will give the same results as the longer form.

The figures and diagrams here given present the results in only one community. As was pointed out in a previous article, results on the Dearborn examinations were not lumped, but were worked up separately. Other results bore out the findings here presented.

Series II was treated in exactly the same way as Series I, and with similar results. It will be unnecessary to reproduce the Table giving the separate test age scores, but Diagram 2 presents part of the evidence graphically. Here we see that the Opposite Completion Test is a very good one for the purposes of the examination. It has a possible score of 34 points, while the 16-year-old children in school score on the average 26 points only. There is no large bunching of scores at the upper end of the curve, and the test would probably differentiate a year or two farther of the selected age groups in school. On the other hand, the Maze Test, which has a maximum of 25 points, shows a median score of only 7 for the sixteenth year. This test took considerable time to give, yet it turned out to aid very little in differentiation.

The Memory for Digits (Ladders) and the Picture Symbol Tests gave similar results. The median score of the 16-year-olds in the Memory Test was only 7 out of a possible 16, while with the Picture Symbols they got only 8 out of a possible 20.

The dropping of the three tests just discussed so decreased the time that it was not necessary to make further deletions. This made it possible to keep the plan of having two parts to the test, a practice which seems highly desirable, as it is thus possible to sample the individual's work at two different times.

Correlations were made with the scores on each form of the examination grouped in 10-point intervals. The tests had been given to the children from the Grade II through Grade XII in three different communities with groups of 1023, 1430 and 1821 cases. The Pearson Coefficients were 0.99, 0.99, 0.96, which show that there was practically no loss in accuracy from dropping the three tests.

It was also found desirable to make a few slight modifications in the tests which were retained. These changes do not alter the

---

[1] The method of grouping the scores may be in some measure responsible for these high coefficients. Scores on the first edition were grouped by fives (0–4, 5–9) and those of the revised examination by tens (0–9, 10–19). Maximum scores are 100 (300 divided by 3) on old test, and 182 on revised form.

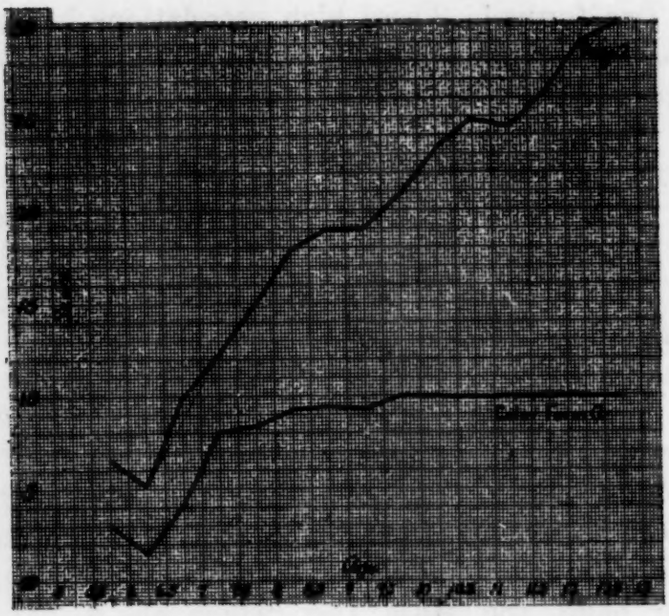

DIAGRAM 1.—Comparison of scores, by half-years, of two Series I Tests.

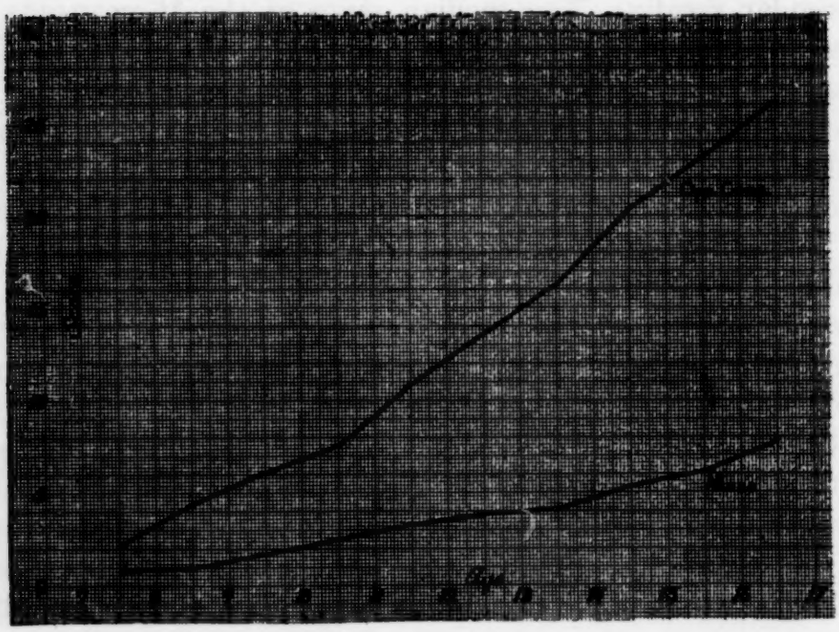

DIAGRAM 2.—Comparison of scores, by age, of two Series II Tests.

tests materially, but aid considerably in the giving of the examination and the correction of the papers. In the First Series a cat has been added in the second item, and directions are now given to draw a dog chasing the cat. A number of "boxes" are added in places where the child is to write numbers. The coins have been enlarged to very nearly their actual size, and the field for the Ball and Field Test has been turned about a bit.

There have been a number of objections made to the tests on pages 1 and 2 of Series 1 because there are no definite time limits for them, and because in some instances the judgment of the examiner is required in scoring. In spite of these objections no change was made in this part of the test. As the two pages in questions are filled with relatively short questions the time limits would necessarily be very short, and the examiner would be forced to pay more attention to his watch than to the group. A loss of spontaneity would result which would be highly undesirable, and almost disastrous in the lower grades. The accuracy of measurements obtained under such conditions would be highly questionable. Another point to be held in mind is the fact that the tests on these two pages are largely adapted from the Binet Tests which impose no time limits, but allow the subject to work at his own speed. Many of the performances upon which the judgment of the examiner is required, such as copying the diamond or square and drawing the star are modeled after the Binet where practically the same method of scoring has been used with much uniformity. Furthermore, there have appeared recently some rather weighty criticisms of examinations which are composed entirely of "objective" tests.[1] For these reasons the writers feel that the retention of the tests in their present form is justified.

The second part of the clock question, in which the children were asked to put hands in clocks to show various times, had given considerable trouble. In responding to the direction: "Show how the clock looks when school begins in the morning," the children were likely to record the time when the school doors opened, the time when they could get into their rooms, the time when they were expected to be in their seats, or some other such time. Also, the sunrise and sunset questions were hard to score. To obviate these difficulties the first three questions call now for definite times; six o'clock, half past four,

---

[1] See, for example, Roback, A. A.: Subjective Tests vs. Objective Tests, *Journal of Educational Psychology*, Vol. XII, No. 8, Nov., 1921.

and quarter of two. Centers have been added to the clock faces to aid the children in their drawing.

Test 17, the First Substitution Test, is given a slightly longer time in the revision. There were no perfect scores on this test, even among the eighth grade pupils, and it was thought that another half minute would produce greater differentiation among the superior pupils. This was desirable because of the omission of the map question, which had been the chief aid in differentiating the superior pupils.

The Picture Completion Test (No. 19) in the first edition consisted of fifteen items and covered three pages. It was found that these items fell into groups according to their difficulty, and for the revision one or two items were picked from each group. This gave a test of seven items, and the size of the pictures was slightly reduced so that all could be placed on a single page. Two practice problems were included, partly to compensate for the practice effect which was present in the longer test, and which probably aided the pupils in the solution of the later items. In order to keep the scores comparable, two points were assigned to each item in the revised form of the test.

Although the map questions are omitted from the new form, the map is reproduced for use in the Ruler Test.

The retained Series II Tests also underwent slight modifications. In Test 1, the Picture Sequences, it had been found that a few of the pictures were very often misinterpreted, and these were changed. The pages have also been improved by putting the practice pictures on the previous page, thus leaving room for blank spaces between adjacent sets of pictures. The time of this test was reduced from 7 to 6 minutes, as it appeared that the shorter time gave ample opportunity for all the pupils to do all that they were capable of doing correctly.

In the Word Sequence Test, also, there were a few items which were somewhat ambiguous, and these were changed for clearer forms. The time limit of this test was also reduced from 7 to 6 minutes.

The Formboard Test had one of the figures out of proportion, and this was redrawn correctly. Also, the figures which are used for demonstration were lettered so that it would be easier for the examiner to refer to them.

No change was made in the Opposite Completion Test (No. 4) except that the time limit was lowered to 7 minutes.

The time necessary for the Proverbs Test was cut down by requiring only the first word to be numbered. A second example was added

TABLE I.—DEARBORN INTELLIGENCE EXAMINATION, SERIES I

Median Scores for Separate Tests by Half Years of Chronological Age.

Test	5–6	6–0	6–6	7–0	7–6	8–0	8–6	9–0	9–6	10–0	10–6	11–0	11–6	12–0	12–6	Maximum score
Page 1, G1	11.0	10.9	13.0	14.7	17.4	18.4	19.1	19.7	19.9	20.7	20.5	20.3	20.8	21.1	21.1	22
Page 2, G1	6.5	5.1	9.9	12.2	14.9	17.8	18.9	19.0	20.8	23.4	24.8	24.5	26.4	29.1	30.0	35
Clocks, G1	0	0	4.1	8.1	10.4	12.4	14.1	17.0	22.3	20.3	22.2	22.4	24+	24+	24+	24
Substitution, G1	0	0	0	0	5.2	7.8	9.1	6.7	10.0	9.8	12.0	11.8	13.3	13.2	14.0	30
Color-form A, G2	3.0	1.5	4.2	8.1	8.4	9.3	9.4	9.3	10+	10+	9.9	10+	10+	10+	9.9	10*
Color-form B, G2	0	1.6	3.8	9.2	13.3	12.6	13.5	13.4	14.9	15+	15+	15+	15+	15+	15+	15*
Digit Memory, G2	0	0	0	6.8	11.0	13.5	13.6	15.0	15.3	16.7	18.2	18.8	19.4	17.5	20.0	24*
Marking Pictures, G2	5.0	5.4	8.3	10.6	10.3	12.2	12.5	12.5	13.0	13.2	13.4	13.7	13.8	13.7	14+	14*
Picture Recognition, G2	2.2	2.4	5.6	9.6	10.6	13.8	13.6	14.5	15.5	16.4	15.6	17.6	17.5	17.6	17.8	20
Color-form C, G2	6.0	3.6	9.1	10.9	11.9	12.2	12.7	12.9	13+	13+	13+	13+	13+	13+	13+	13*
Dominoes, G2	1.8	3.4	4.4	6.1	7.3	6.9	8.0	9.4	10.0	9.7	10.3	10.7	10.9	10.6	10.9	13
Picture Completion, G3	4.0	4.4	4.1	5.3	8.0	10.6	10.9	12.3	13.1	13.0	13.8	14.4	14.3	14.0	14.3	15
Map, G3	0	0	0	4.6	7.1	7.0	9.5	10.5	10.4	11.7	11.2	14.0	14.9	15.2	15.2	20*
Ruler, G3	0	0	1.2	1.9	4.5	4.3	7.0	6.7	7.8	7.6	9.1	10.5	12.2	12.1	12.4	20
Substitution, G3	0	0	1.9	7.0	12.5	16.0	15.0	15.0	16.5	17.0	17.0	17.6	20.4	18.8	19.6	25
Number Cases	32	55	58	51	69	53	55	60	62	49	66	61	54	59	48	

The ages at the head of the columns indicate the lower limits of the groups. Thus the first column shows the median scores of all pupils from 5–6 to 5–11 inclusive.

The plus sign after a median indicates that more than half the pupils in the group made a perfect score in that test.

* shows tests which were eliminated.

for demonstration, and one or two changes were made in the tests at points which had given difficulty. There was some thought of dropping out the part of this test in which the children are asked to explain proverbs in their own words as the scoring of this part of the test cannot be made as objective as the rest of the tests. However, this part of the test proved excellent for picking out the very superior individuals, so it was finally retained.

A number of the drawings in the Faulty Pictures Test proved very unsatisfactory, and so the upper half of this page was redrawn. No change was made in the time limit.

In the first part of the Number Puzzles a correction was made in the perspective of the last diagram. The time for this test was reduced to 2 minutes, as it was found that practically everyone, even in the lower grades, could finish within this time. In the second part the demonstration examples were changed by leaving out the sums on the columns and rows, and a figure was omitted from the first of these. The two problems which had more than one possible solution were changed.

These revisions have made of both series considerably more convenient instruments for both tester and scorer. There is little loss of the range of variability, except among the very superior pupils in Series II, and discrimination at the extreme upper end of the scale is not ordinarily of great practical importance. Very convincing evidence will be presented in a forthcoming article that in these tests the average adult age is approximately 14½ years. The differentiation of superior adults is made up to intelligence quotients of 140. Further differentiation of the small number of cases beyond this level would ordinarily be made by individual examinations.

The saving of time is especially valuable in Series II, as either part of the revised test can be given in the ordinary high school period. It is possible, when a rapid survey is necessary, to use Examination A alone in the first three grades, and Examination C in the others. The correlations of these separate parts with the whole examinations are high, running well over 0.90 in a number of communities. This practice is not recommended, however, as the more tests we have the more likely we are to avoid the contingency that a chance failure or success in any one test might misrepresent the individual's true capacity.

NOTES ON ARTICLES IN EDUCATIONAL PSYCHOLOGY IN CURRENT ISSUES OF OTHER MAGAZINES

REPORTED BY CECILE COLLOTON

Department of Educational Psychology, The Lincoln School of Teachers College

Intelligence Tests

Bibliography of Tests for Use in High Schools. Edited by C. C. Tillinghast. Teachers' College Record, 1922, September. Annotated list of intelligence and educational tests for use in high schools.

The Contribution of Intelligence Tests to Educational Guidance in High School. I. N. Madsen. The School Review, 1922, November, 692–701. Shows relation between level of intelligence and kind of high school work a student may profitably do. Nine tables give detailed data.

Adult Tests of the Stanford Revision Applied to University Faculty Members. Helen H. Caldwell. Journal of Experimental Psychology, 1922, August, 247–262. Results of the Binet Test of 34 college professors analyzed. Conclusions as to comparative ability of literary and scientific groups.

Intelligence—a Definition. Lightner Witmer. The Psychological Clinic, 1922, May–June, 65–67. Another discussion as to what intelligence really is.

Slot Maze A. Herman H. Young. The Psychological Clinic, 1922, May–June, 73–82. Description of a slot maze, the first of a series now being devised and standardized. Results of the use of the maze in testing 1304 children, from 4 to 9 years of age.

The Present Status of Mental Testing. Stephen S. Colvin. Educational Review, 1922, November. The uses of intelligence tests classified and described. Conclusions as to the place of intelligence testing in the present scheme of education.

Classification by Test of Intelligence of Ten Thousand First Grade Pupils. Charles S. Berry. Journal of Educational Research, 1922, October, 185–203. Classification into three groups by the Detroit First Grade Intelligence Test. A study of the results.

The Relation between College Failures and Mental Tests. Franklin O. Smith. School and Society, 1922, Oct. 14, 444–446. Using the Thorndike Intelligence Test and Smith's Graded Intelligence Test for High Schools and Colleges to determine fitness for college work. A study of 306 freshmen in the State University of Montana.

Intelligence Tests as a Means of Admitting Special Students to Colleges and Universities. William M. Proctor. School and Society, 1922, October 21, 471–476. Records of 41 Veterans' Bureau trainees at Stanford University show success in university work in direct ratio to scores on the Army Alpha. Entrance requirements should be waived in special cases.

The Competency of Fifty College Students. Karl G. Miller. The Psychological Clinic, 1922, March–April, 1–25. Study of data collected in a class in elementary psychology at the University of Pennsylvania. Analysis of the 50 individual records.

The Correlation between the IQ and the AQ. Andrew H. MacPhail. School and Society, 1922, Nov. 18, 586–588. A summary of the results of projects carried out under Professor Colvin at Brown University. Correlation from 0 to −0.71 between IQ and AQ are reported. The IQ is derived from Group Tests.

Validating Intelligence Tests. Garry C. Myers. School and Society, 1922, Nov. 25, 612–614. Shows the importance of adequate criteria and sufficient data in attempting to validate tests.

Are Mentality Tests on Right Lines? William R. Comings. Educational Review, 1922, December, 392–394. Pleads for a comprehensive examination of children with mental age as only one element. Vital question to be answered is, "What can each individual do and what can he do well?"

The Intelligence Quotient From Two Viewpoints. W. T. Root. Journal of Applied Psychology, 1922, September, 267–275. Takes the position that in dealing with the masses the IQ is undoubtedly constant but that there are likely to be individual cases of fluctuation due to conspicuous changes in one or many of the factors producing intelligent behavior.

An Experiment in Classifying Primary Grade Children by Mental Age. Frances Lowell. Journal of Applied Psychology, 1922, September, 276–290. Details of the reorganization of the first, second, and third grades in the City Normal School of Rochester, N. Y., on the basis of mental age. Need for special curriculum for slow, dull pupils. Better promotion results than by chronological age classification.

Hormones and Intelligence. Sante Naccarati and R. L. Lewy-Guinzberg. Journal of Applied Psychology, 1922, September, 221–234. Correlation between the ratio of height to weight and intelligence in normal individuals. Tables are given showing all measurements on 396 cases.

Intelligence Quotients of Adults and Related Problems. Walter F. Dearborn. Journal of Educational Research, 1922, November, 307–325. Presents evidence to show that the average adult mental age is approximately 14½ years and advocates the calculation of the IQ on this basis. An appendix gives a "method of translating adult test scores into intelligence quotients."

A Mental Alertness Test. A. W. Kornhauser. The Journal of Personnel Research, 1922, July, 103–121. Practicability and desirability of using intelligence tests in certain phases of personnel work. Brief illustration of the test with illustrative examples.

Interpretations of the Reactions of the Feeble-minded on the Healy-pictorial Completion Test II—Social Implications. Doris E. Perry. The Journal of Delinquency, 1922, March, 75–85. IQ not adequate for discovering feeble-minded having traits most necessary for social adjustments. Analysis of test results and social behavior of irregular cases.

Objective Mental Diagnosis. E. A. Doll. Journal of Delinquency, 1922, May, 119–131. Description of the various types of tests used in the New Jersey State Home for Boys. Illustrations of reports used.

Notes on Racial Differences. A. M. Jordan. School and Society, 1922, Oct. 28, 503–504. Data collected at Fort Smith, Arkansas show pronounced differences

of attainment of whites and of negroes between the ages of 10 and 14 in the National Intelligence Tests.

The Influence of Intelligence on the Selection of Associates. John C. Almack. School and Society, 1922, Nov. 4, 529–530. An expression of 387 children in grades 4 to 7 inclusive of the public schools of San Jose, California respecting work and play associates shows tendency for an individual to choose his associates from his own mental level.

Educational Tests

A Criticism of Tests Requiring Alternative Responses. H. H. Hahn. Journal of Educational Research, 1922, October, 235–240. Objects to "yes-no, same-opposite, true-false" tests on grounds of inaccuracy and reactionary influence on teaching.

A Study in Correlation of Allied English Abilities. Grace S. Jamison. Journal of Educational Research, 1922, Oct., 241–253. A study of the relation between language ability and written composition. Achievement measured by Trabue Completion Tests and the Nassau Supplement to the Hillegas Scale. Four tables give correlation results.

A True-false Test in Literature. Arthur G. Melvin. The English Journal, 1922, October, 491–496. An adaptation of the true-false type of examination to English. The examination is given in full with the scoring key and the results.

The Barr Diagnostic Tests in American History. C. W. Odell. School and Society, 1922, Oct. 28, 501–503. Results of the use of the Barr tests in the American History Classes of the South Bend, Indiana, High School. Correlations of the subtests with each other and the total score are given.

Comments on the Plan for Sectioning Classes on the Basis of Ability. C. E. Seashore. School and Society, 1922, Nov. 4, 514–517. Selected extracts from replies to an inquiry of a large number of colleges and universities as to the extent to which sectioning of classes on the basis of ability is in effect or under consideration.

Miscellaneous

Recognition of Individual Differences in the Junior High School. August Dvorak. School Review, 1922, November, 679–685. An analysis of the results of a questionnaire study of the practices of 86 junior high schools in various parts of the United States. Individual differences recognized but not adequately provided for.

Adjustments in California. Will C. Wood. Journal of Educational Research, 1922, November, 326–331. Big problems in educational research that educational testers can help to solve.

Statistical and Non-statistical Interpretation of Test Results. Samuel W. Fernberger. The Psychological Clinic, 1922, May–June, 68–72. Diagnostic value of tests more important than final test scores. Non-mathematical interpretation of results more significant than statistics.

More Accurate Use of Composition Scales. Edward William Dolch. The English Journal, 1922, November, 536–544. Points out three distinct sources of inaccuracy in the use of English composition scales and suggests ways of meeting these difficulties.

Using the Results of Measurements in Reading in Training Student-teachers. M. Elizabeth James. The Elementary School Journal, 1922, November, 190–196. Original exercises in checking comprehension and improving rate in reading. Remedial measures used in individual cases described.

The Conversion of Test Scores into Series which Shall Have Any Assigned Mean and Degree of Dispersion. Clark L. Hull. The Journal of Applied Psychology, 1922, September, 298–300. Describes a formula for making scores from a number of tests directly comparable.

A Method For the Study of Vocational Interests. Max Freyd. Journal of Applied Psychology, 1922, September, 243–254. A detailed description of two questionnaires for studying the interests, likes and dislikes of an individual as a dynamic basis for vocational guidance and selection.

A Comparison of the Achievement of High School and University Students in Certain Tasks in Chemistry. S. R. Powers. Journal of Educational Research, 1922, November, 332–343. Students of chemistry in the larger high schools do practically as good work on this particular test as university freshmen.

The Homewood Demonstration School at Johns Hopkins University. M. Rose Patterson, School and Society, 1922, Nov. 18, 577–584. Description of the conditions under which a half year's work was covered successfully in 2 months.

A Comparison Test for Investigating the Ideational Content of the Moral Concepts. R. A. Brotemarkle. Journal of Applied Psychology, 1922, September, 235–242. Description of a test designed to measure the background of "moral concepts" of an individual. Results of the tests of 331 college students compared with their quintile ranking on other tests.

The Psychology and Pathology of Personality. Vernon M. Cady. Journal of Delinquency, 1922, September, 225–232. "A summary of test problems and bibliography of general literature."

When Children Read for Fun. Jenny R. Green. School and Society, 1922, Nov. 25, 614–616. Reports an experiment with 500 sixth and seventh grade children to find correlation between reading ability and material chosen when "reading for fun."

An Ancient Score Card. Elmer E. Jones. The Psychological Clinic, 1922, March–April, 26–35. Description of a score card used 20 years ago. Its predictive value as shown by the later achievement of individuals whose cards are reproduced.

Diagnostic Teaching. M. Alice Weir. The Psychological Clinic, 1922, May–June, 116–122. A case study.

A Study in Grades and Grading under a Military System. Robert L. Bates. Journal of Experimental Psychology, 1922, October, 329–337. Describes the rigid system of grading in use at the Virginia Military Institute, Lexington, Va.

An Ethical Discrimination Test. S. C. Kohs. Journal of Delinquency, 1922, January, 1–16. A new battery of tests to measure other qualities than intelligence. Two tests are new—others are new compilations of old material. Not yet standardized. May be used either as individual or group test.

Individual Variability in Test Performance. Mervin A. Durea. Journal of Delinquency, 1922, March, 86–98. Detailed study of 150 cases—50 at each of the mental age levels, 9, 10 and 11. Comparisons of four types of tests—Stanford-Binet, Performance Tests, Association, and Literacy.

Home Conditions and Native Intelligence. W. W. Clark, Journal of Delin-

quency, 1922, January, 17–23. The Whittier Scale used for grading home conditions of boys committed to state industrial school. Four tables compare general quality of homes and IQ's of boys.

The Functions of Tests and Scales as Found in Recent Educational Literature. A. G. Capps. Journal of Educational Research, 1922, October, 204–208. Reports an investigation of 50 different educational books and magazine articles. Detailed tabulations and summaries of the topics noted in the title are given.

The Theory of Differential Education as Applied to Handicapped Pupils in the Elementary Grades. J. E. Wallace Wallin. Journal of Educational Research, 1922, October, 209–224. Recommends an organization of special classes for deficient or defective pupils. Illustrates the application of such class organization. Pleads for trained workers and a careful use of tests.

Calculation and Interpretation of Percentile Ranks. L. L. Thurstone. Journal of Educational Research, 1922, October, 225–235. Description and explanation of the percentile curve and its graphical calculation.

The Age-grade Status as an Index of School Achievement. Charles L. Harlan. Educational Administration and Supervision, 1922, October, 413–423. Mental ability as shown by intelligence tests and school achievement as shown by educational tests better bases for classification than age or years in school. Eight tables give detailed data.

Another Educational Campaign. J. W. Richardson. Journal of Educational Research, 1922, September, 97–101. An experiment in the teaching of spelling and its results.

The Variability of Teachers' Marks. Nathan Silberstein. The English Journal, 1922, September, 414–424. An analysis of the rating of seven questions on a Regents examination paper by 31 members of an English department.

The Psychology of Solving Puzzle Problems. Joyce E. Mather and Linus W. Kline. The Pedagogical Seminary, 1922, September, 269–282. An experimental study of reasoning, using mechanical puzzles as material. Full details as to procedure and results.

The Reliability of Judgment of Personal Traits. John Slawson. Journal of Applied Psychology, 1922, June, 161–171. Twenty-five teachers in each of six schools were given independent ratings in 11 personal traits by four or more judges. Results are studied in detail.

Computing Intercorrelations of Tests on the Adding Machine. Herbert A. Toops. Journal of Applied Psychology, 1922, June, 172–184. Details of a procedure for the rapid calculation of intercorrelations.

Ranking Students from Their Literal Grades. J. Charles Rathbun. School and Society, 1922, Sept. 16, 326–335. Proposes a method of evaluating grades on the basis of the normal probability curve.

Standardization of the Whipple-Healy Tapping Test. Amy Hewes. Journal of Applied Psychology, 1922, June, 113–119. Describes in detail the establishing of age standards for a test measuring speed and accuracy of eye-hand coordinations.

The Reconstruction of Democracy. George B. Cutten. School and Society, 1922, Oct. 28, 477–489. The value of an intellectual aristocracy. The importance of the general and normal training of the "mentally fittest."

Facing the Facts. W. L. Ettinger. School and Society, 1922, Nov. 4, 505–512. Describes the New York school situation with regard to ages of pupils, rates of progress, achievements of class groups, etc. Need for better classification.

NEW PUBLICATIONS IN EDUCATIONAL PSYCHOLOGY AND RELATED FIELDS OF EDUCATION

1. *A Textbook of Psychology.*—The author[1] of this book emphasizes its scientific character even to the extent of including the term "scientific" in the title. Again in the preface, we are told that this book represents a definite departure from tradition. Just what tradition is meant, the reviewer finds it difficult to imagine. At any rate, the book itself does not differ from the conventional textbook of psychology in any material way. We find the usual lengthy treatment of the senses, sense perception and the nervous system. Indeed, fully two-thirds of the book is devoted to these topics. Scattered throughout the pages, we find several instances of our present-day popular psychological game of baiting the introspectionist, which to the reviewer, seem out of place in an elementary text.

The author, however, is no behaviorist, since he describes psychology as "a science of the conscious responses of the organism." He dwells upon the value of psychology in furnishing materials applicable to the problems of physical science, education, industry and the arts, and to social problems. Little attempt throughout the book is made to show the student the applications of the various topics discussed to the above-mentioned fields.

In the chapter on thought and thought content, there is much discussion of "imaginative types," such as "the visile, audile, tactile, gustile, olfactile and motile." Because "workers in mental measurement are for the most part persons of intensely practical bent," they have neglected the study of imaginative types. To discover the various types, we must use the method of simple report, which begins "Imagine a pan of onions frying on a stove." And then follow the usual questions.

The book is interesting, and the discussion of sensation and perception is detailed and well done. It will not prove an easy book for the beginner. For the student of education or the educational psychologist, it has nothing of particular interest.

R. P.

---

[1] Dunlap, K.: "The Elements of Scientific Psychology." Mosby, St. Louis, 1922, pp. 368.

2. *A Book about Tests.*—This is the age of standard tests, educational and intelligence. Teachers, superintendents and others are using them by the million, and are finding them instructive and useful for all sorts of purposes. The number of books about tests is very small and a new one is, therefore, welcome. The authors of the present volume[1] have tried to explain in simple terms what standard tests are and for what purposes they should be used. They devote a chapter to a description and discussion of the chief tests in each subject, such as arithmetic, reading, and the like. Two chapters are devoted to intelligence tests and naturally the treatment has to be brief and sketchy. A glossary of terms at the end of the book should be useful to the beginner in the subject.

The book commends itself to the reviewer for its simple and pleasant style. It should prove extremely useful to the thousands of teachers who are anxious to know what this testing movement is all about and who have not the time or inclination to delve deep into the subject.

R. P.

---

3. *The Recorded Activities of Six-year-olds in an Experimental School.*[2]—An unusual conception of the nature and method of curriculum reconstruction is exemplified in this booklet. After questioning the fundamental assumptions underlying current practices and points of view, these workers have endeavored to blaze a new trail. The significant fact in this connection is that they have kept a careful day by day record of the resulting activities, and thus made available a mass of data necessary to the evaluation of their work.

The necessary materials and furnishings are listed and a typical program is presented. Then the record of each week's work is given under the following heads: (I) Play experiences, (II) Practical activities, (III) Special training, and (IV) Organization of information. Under the third heading the reader is surprised to note the emphasis on number experiences, the formal nature of some of the special training in music and the evident predetermined postponement of reading. The notes on language work show that the experimenters

---

[1] Pressey, S. L. and Pressey, L. C.: "Introduction to the Use of Standard Tests. A Brief Manual in the Use of Tests of Both Ability and Achievement in the School Subjects." World Book Company, Yonkers, N. Y., 1922.

[2] Scott, Leila V., Pratt, Caroline, and others: *Record of Group VI*, 1921. City and Country School, New York, 165 West 12th Street, pp. 72.

have formulated with more or less clarity a hypothesis concerning the approach to reading. We can only hope that the conditions of the whole experiment may be controlled and set with sufficient care to test that hypothesis scientifically.

There is no clear statement of objectives and the reader is left to evaluate the "curriculum" as he sees fit. New criteria are necessary to determine the validity of new departures and those who initiate untried procedures should formulate the objectives in the light of which a true critique of their endeavors is possible.

Those who consider the record a proposal miss the point. The significance of adequate records in furnishing points of reference and departure in experimentation is seldom realized because they are infrequently put into shape and published. The technique of record keeping needs illustration. The analysis of the group records and study of the growth of individuals is a further step of experimental technique, facilitated, but not exemplified by this booklet.

L. Z.

---

4. *A Manual of Suggestions on Education Related to Sex.*[1]—Prepared under the direction of the Surgeon General of the United States Public Health Service in collaboration with the Bureau of Education, this Manual is well sponsored. There are those whose prejudices will incline them to dismiss this book without a hearing. It would be salutary, indeed, if such persons could be induced to read the list of 60 contributors whose names appear in the preface, and enough of the introduction to enable them to realize how much of their social and civic responsibility they are side-stepping by a closed-minded adherence to a position which has become untenable. Because most conscientious objectors have considered the subject too narrowly, we are prompted to quote this definition of the problem:

"In harmony with the rapidly developing psychology of education, it becomes necessary to conceive of education in relation to sex as but a phase of character education as a whole. As such, 'sex education' means vastly more than instruction concerning sex; it means a comprehensive and progressive process of care, guidance, and example extending over a long period of years, from infancy to maturity. Moreover, sex education is a social and a socializing process; both in its progress and in its results it reaches far beyond the boundaries of the physical person. Because of the far-reaching effects of the eventual attitude and practices of the

[1] Gruenberg, Benjamin C., Ph. D., Editor: "High Schools and Sex Education." Government Printing Office, Washington, 1922, pp. VII + 98. $0.50.

individual, sex education carries with it obligations of the widest social importance. As a phase of character formation, sex education must include all the instruction and training that may help to form normal and wholesome attitudes and ideals in relation to sex, and to shape conduct in accord with such attitudes and ideals. Such education must, therefore, be developed as an organic part of the entire educational program, and not be considered a special and isolated bit of ritual to be performed at a given time, and then dismissed as finished."

The Manual incorporates the results of experimentation for the practical guidance of thousands of teachers who realize the significance of their tasks but are not equipped to proceed without guidance.

Three chapters are devoted to the general aspects of the problem. Each of the succeeding seven chapters then takes up the relation of sex education to the content of the course in some subject. To illustrate we quote the list of topics treated in Chapter Ten, under the heading, The English Course: "(1) English and literature, (2) Literature and life, (3) Teaching discrimination, (4) Positive ideals, (5) The negative side, (6) Literary biography, (7) English composition, (8) Supplementary reading. References." What is narrowly conceived as sex education is treated in Appendix A under emergency measures. The appendix also offers a suggested outline of a summer school course for teachers, a form for a physical examination record card, and a classified bibliography.

Those who realize the psychological implications of the problem must be encouraged by the insight revealed in the organization of the materials contained in this manual.

L. Z.

---

5. *A Case Book for Teachers and Parents.*—This is the subtitle of a volume[1] in which the author has drawn on his experience for a large number of instances illustrating the numerous problems which confront those who deal with adolescents. The work is at once a plea for a more sympathetic treatment of youth and a compendium of practical suggestions. While the material is very loosely organized and the style is extremely discursive, the author will hardly be considered at fault by those who are disinclined to consult more scientific treatises. The discussion is limited to the description and disposition of 'cases' as they arise. The evidence is not summarized to point out

---

[1] Stableton, J. K.: Your Problems and Mine in the Guidance of Youth. Public School Publishing Co., Bloomington, Ill., 1922, pp. IX +274.

the need for a fundamental change in educational policies and provisions. Nevertheless, the book will be useful to those who must find in vicarious experience first aid for pressing problems.

L. Z.

---

6. *A Bulletin on Educational Measurements.*—The program of the Ninth Indiana University Conference on Educational Measurements furnishes the content of an interesting bulletin[1] which contains, among other things, four short articles on educational provision for exceptional children by Henry H. Goddard and three brief papers on reading by William S. Gray. Professor W. F. Book and two research workers at Indiana University report on the formulation and administration of "Group Tests for Measuring Observational Learning and Accuracy of Report," and present experimental evidence on the educability of the traits in question and their relation to intelligence and school training.

The reader will be surprised to find Prof. W. M. Black's searching philosophical discussion of "Three Remedial Principles in Education" in the same bulletin.

L. Z.

---

7. *The Relation of Speed and Accuracy of Performance.*[2]—In judging the lengths of lines, the magnitude of weights and the quality of handwriting, Garrett found that increasing the rate of judgment from one every four seconds to one in two seconds brought about a gradual increase in accuracy which could not be maintained under greater speed—one or one-half second intervals. In simple motor acts, such as thrusting and tracing, an inverse relation between speed and accuracy was apparent. The author concludes, from the experiments upon judgment at least, that for each individual there is an optimum rate which permits most appropriate readjustment between reactions. The optimum interval catches the subject at the peak of the mental integration. Whether the optimum rate could be appreciably increased by practice was not disclosed with certainty.

A. I. G.

---

[1] *Ninth Conference on Educational Measurements.* Indiana University Bulletin, The Extension Division of Indiana University, Bloomington Vol. VII, No. 12, 1922, pp. 146

[2] Garrett, H. E.: A Study of the Relation of Accuracy to Speed. *Archives of Psychology,* Columbia University, New York, No. 56, 1922, pp. 104.

8. *Intelligence and Tonsils.*[1]—Using the control group technique, the influence of the removal of diseased tonsils in the case of 236 children was determined by repeated examinations before and after for a period of a year or more. An increase in weight—slight during the first but greater during the second half year following the operation—was the only trait in which the "tonsil" group exceeded the control group. The removal of tonsils has no effect on the growth of height, strength of grip, speed of tapping, or in performance in the Healy or Stanford-Binet Intelligence Tests; at least, not within the year following removal. The author accounts for the improvement in school work in terms of increased health, vigor, and "volitional and emotional normality."

A. I. G.

---

9. *The Effect of Mood upon Performance.*[2]—The student subjects, previous to a series of tests of mental and motor functions, indicated their temporary mood on a rating scale which embraced steps from the "least" to the "greatest cheerfulness ever felt." Subjective estimates and tests were repeated daily for a month or more. With improvement from practice eliminated, correlations between performance and cheerfulness were computed for each subject. The coefficients clustered about a central tendency of approximately zero, *i.e.*, the influence of mood upon achievement was scarcely appreciable. There was, furthermore, no substantial evidence that some individuals, much more than others, are functionally impaired by the "blues." Mood, insofar as this investigation goes, falls in the same category with "feelings of fatigue," unusual temperature and humidity, noise and other influences whose effect on achievement, often subjectively estimated as great, turns out to be insignificant when objectively determined.

A. I. G.

---

[1] Rogers, Margaret Cobb.: Adenoids and Diseased Tonsils: Their Effect on General Intelligence. *Archives of Psychology*, Columbia University, New York, No. 50, 1922, pp. 70.

[2] Sullivan, E. T.: Mood in Relation to Performance, *Archives of Psychology*, Columbia University, New York, No. 53, 1922, pp. 71.

10. *The Influence of Incentive and Punishment upon Reaction-Time.*[1]—Measurements of the speed of motor reaction to an auditory stimulus were recorded under three conditions: (1) "Normal," in which only general directions for procedure were given; (2) "incentive," in which the record for the preceding trial was disclosed; and (3) "punishment," in which the subject, by means of a mechanical device was given an electric shock for relatively slow reactions. The average saving due to "incentive" was 6 per cent; that due to "punishment" 15 per cent of the "normal" time. In terms of sigma (thousandth of a second) the amounts of saving were 8 and 20, respectively.

A. I. G.

---

11. *Psychology Classics.*—Professor Knight Dunlap has conceived a splendid idea in his plan to reprint older books or articles that have been of significance in the history of psychology. To this proposed series he gives the name "Psychology Classics" and Volume I of this series has just appeared.[2] This contains a reprint of James' article first published in Mind (1884) and of his chapter on the emotions in his Principles of Psychology (1893). Along with these two selections we have also a new and very happy translation of Lange's *Ueber Gemuethsbewegungen,* made by Miss Istar A. Haupt from Kurella's German version. According to the editor this is the first English translation of Lange's important work. We thus have in one handy volume the original statement of what is now generally known as the James-Lange Theory of the Emotions. The volume will be appreciated by all students of the emotions and of the history of psychology.

R. P.

---

12. *How One Superintendent Made Use of Standard Tests.*—In this book[3] we have the story of how a district superintendent made use of standardized tests in his schools and a very interesting story it is.

---

[1] Johanson, Albert M.: The Influence of Incentive and Punishment upon Reaction Time. *Archives of Psychology,* Columbia University, New York, No. 54, 1922, pp. 53.

[2] Lange, C. G. and James W.: The Emotions. *Psychology Classics,* Vol. I. Williams and Wilkins, Baltimore, 1922, p. 135.

[3] Brooks, S. S.: "Improving Schools by Standardized Tests." Houghton, Mifflin Company, 1922, p. 278.

As the author himself says, "It is the story of how a corps of faithful, hard-working, but mostly untrained teachers, with the aid of an inexperienced superintendent, put standardized tests and measurements to practical use throughout a school system to the considerable advantage of all concerned." The author shows us step by step just what he did, and not the least interesting is the ingenious way in which he brought his teachers to see the need of tests and to feel that the results were something that would be of help to them.

Most of the tests used were educational tests, and the progress of pupils was measured. In the schools, as a whole, there was a tendency to rank decidedly above grade in arithmetic. This the author believes was due to the large amount of time spent on this subject, and it was decided to cut down the time devoted to arithmetic. On the other hand, the reading scores were "scandalously low," and there seemed to be many reasons for this. The superintendent and teachers, therefore, concentrated on the problem of reading. Several chapters in the book deal with the topic of reading and indicate what methods were employed to improve in this subject. And, lastly, we are told what was done to help the children in their methods of study.

The book is not a text on tests or measurement or reading or methods of study. A casual glance at the book makes one feel that it is a mixture of all these things. And so it is, and for a very good reason. Standard tests showed certain deficiencies in the school system, and in this particular case the results of the tests made superintendent and teachers concentrate on the problems of reading and children's methods of study. In other school systems, tests will reveal other deficiencies and other problems for study. And this is the real justification for the use of tests, namely, that they lead on to further problems and make more efficient teachers and pupils. Every superintendent ought to read Mr. Brook's book, not for the purpose of doing just what he did, but as a suggestion as to what problems can be attacked by means of educational and intelligence tests.

R. P.

---

13. *A Scientific Analysis of the Various Types of Silent Reading.*[1]—Those who are familiar with earlier reading studies from the University

[1] Judd, Charles Hubbard and Buswell, Guy Thomas: Silent Reading: A Study of Various Types. *Supplementary Educational Monographs*, No. 23, Department of Education, University of Chicago, Chicago, 1922, xiii + 160.

of Chicago laboratories will welcome the announcement of another monograph, the seventh of an illustrious series dealing with some phase of that subject. The purpose of this study is to ascertain the fundamental differences in the mental processes involved in various types of silent reading and to contribute scientific data from which the somewhat obscure symptoms of success and failure may be ascertained. The substitution of objective evidence for unanalyzed assumptions should lead to revised conceptions of reading ability, and the practical implications should affect instruction in every subject in grades above the third.

Ninety plates and 21 tables are presented and the reader can discern the behavior of the eyes of a number of selected subjects in a variety of silent reading activities. The materials used include paragraphs of the Gray Oral Test, simple prose fiction and excerpts from poems of various degrees of difficulty, passages from geography, rhetoric and algebra texts, French prose and selections from a French grammar, Latin prose and selections embodying two languages. The subjects read and sometimes reread for various purposes, and the records enable the reader to study the effects of changes in purpose and attention, as well as adjustments necessitated by differences in content, difficulty and language. We quote from the monograph. "Whenever the mental processes of pupils show fundamental differences, practical school procedure will have to fit its methods to these differences. The program of co-operation between science and practical teaching is easy to lay out when we thus see the intimate relation between methods and psychological distinctions. The duty of the scientist is to devise methods of discovering and describing fundamental distinctions. The duty of the teacher is to develop practical ways of dealing with the various kinds of mental proccesses which are pointed out."

Some sweeping assumptions of uncritical psychological thought are refuted:

"There is a popular psychology which assumes that, when the reader sees a word, the interpretation is tied by what older psychologists called an association and by what some recent writers have called a bond, to the impression received. The notion that a word and its meaning are two fixed pieces of experience which can be tied together, is a purely mechanical theory and not adequate as a basis for real teaching. The simple mechanical explanation is very satisfying to some minds because it reduces all teaching to the same formula. One can teach little children and big children the same way. All one

has to do is to set up more and more combinations (bonds) and strengthen them. This type of popular psychology finds very little support in analytical studies such as the present monograph reports. A printed page turns out to be—a source of a mass of impressions which the active mind begins to organize and arrange with reference to some pattern which it is trained to work out. The business of the student of educational science is to discover the various ways in which impressions can be organized. Thus the pattern or attitude determines the process. *Mental life is a complex of organized attitudes —not a collection of mechanical associations or bonds.*" This pregnant sentence was hidden in a paragraph and is here given the emphasis warranted by its significance. Such fundamental considerations are presented in Chapter I. The next three chapters give the data. The possible reactions to difficulties of vocabulary, sentence structure and logic are recorded in Chapter II. Chapter III shows that the effects of changes in attention are not the same for all subjects. Some pupils seemed to be unable to control or change the level of attention when asked to study hard, read carefully, skim or prepare to repeat verbatim. Most of the records show that effort results in a narrowing of the span of recognition, lengthening of the pause or fixation, movements of the eye back over the line (regressive movements). Some records show that the subject evades difficulties in silent reading by passing over them without a feeling of responsibility.

Chapter IV deals with analytical study. The plates show that the mental tension and the process involved are very different from those involved in reading, and that the grade-placement and time allowance for such study may be extermined experimentally.

Chapter V shows conclusively that Latin students do not actually read, but fumble and puzzle out meanings in a highly ineffective way. The number of fixations to the line of Latin prose is astounding. The utter helplessness of the Latin student debarred from reference to his vocabulary reminds one of the futile pacings of caged animals. The influence of the direct method of teaching French may be noted when French and Latin records are compared.

Chapter VI is full of suggestions as to practical implications and the necessity for the analytical determination of the bases instructional policies.

While we had hoped to find reports of the effects of various types of training, and thus evidences of the need and effect of specific instruction, that may not have been possible in this study. High

school people should see the necessity and feasibility of scientific attack on more of their problems from this significant contribution to the literature of educational psychology.

L. Z.

---

14. *Achievement "Norms" for Subnormal Children.*[1]—The results of the first survey of a selected group of subnormal children are presented in a recent bulletin over the signature of Dr. Wallin.

Various degrees of subnormality are represented. The tests used were part of the Ayres Spelling Scale and Starch Spelling Tests, the Gray Oral Reading Test, and the Cleveland Arithmetic Test. The results of individual examination by some form of the Binet Test are also given. The writer considers the data of value in determining the classification and educational treatment of subnormal children elsewhere.

Scores are presented in sufficient detail to permit a variety of careful comparisons and the histories and description of outstanding cases lend concreteness to the discussion.

L. Z.

---

15. *The Reading Process.*[2]—This is the title of a new book which aims to put at the disposal of teachers the fundamental considerations underlying method in the teaching of reading. The content of the book—as well as its arrangement—has been used in teacher-training courses. In the opening chapters the discussion is concerned with the sociological and psychological aspects of reading as an elaborate form of language behavior. The evolution of all written language is traced, and the idiosyncracies of English spelling are shown to have a bearing on problems of instruction.

After a brief historical sketch of methods used in teaching beginners to read, there follows a concise summary of the leading reading investigations. Three chapters cover as many fields of investigation. The discussion of oral and silent reading gives the quantitative experimental evidence upon which recent tendencies are based. The future of

[1] Wallin, J. E. Wallace: The Achievement of Subnormal Children in Standardized Educational Tests. *Bulletin* No. 7, Series XX, Miami University, 1922, p. 97.

[2] Smith, William A.: "The Reading Process," The Macmillan Co., New York, 1922, pp. XII + 267.

school readers is discussed in the light of their evolution. The final chapter is devoted to a survey of standardized tests for measuring reading ability. At the end of each chapter there is appended a selected bibliography. Over a hundred titles are listed and these represent the work of more than sixty contributors. The book contains numerous illustrative plates and tables as well as an index.

L. Z.

---

16. *A Study of the Traits of Chinese Students in America.*—Most of us who have attended an American University have seen the Chinese student. We know him to be a clean, serious, earnest, thorough young man or woman, enthusiastic about the academic activities of the campus, little participating in the social or athletic activities. Have we realized that these Chinese students in our American universities are going back to be leaders in Chinese life, and that they are here obtaining that training which will enable them to achieve that leadership and wield it for the common good? How to make careful selection of the student and plan that his stay here may be most profitable is, therefore, a distinct problem for those with authority. Dr. Chu has made a significant contribution[1] in determining certain of the factors which contribute to the success of Chinese students while in America. He has found that knowledge of English is correlated highly with both scholarship ($r = 0.719$) and leadership ($r = 0.679$); but that knowledge of Chinese correlates much lower with scholarship ($r = 0.389$) and leadership ($r = 0.309$). Knowledge of English and knowledge of Chinese are totally uncorrelated ($r = 0.024$). Dr. Chu says, "Since knowledge of English is correlated rather high with both scholarship and leadership and since knowledge of Chinese is correlated rather low with both of them, it is evident that in the preparation of Chinese students to be sent to America more emphasis should be placed upon knowledge of English and less upon the knowledge of Chinese if these students are expected to do well in scholarship and leadership in America." A correlation, however, does not necessarily imply any such causal relationship. It might be that the higher correlation of English than Chinese obtains through a higher common correlation with a third factor such as intelligence. Dr. Chu partly

---

[1] Chu, Jennings P.: Chinese Students in America: Qualities Associated with their Success, T. C. Contributions to *Education* No. 127.

realizes this when he says, "it is very likely and safe to say that intelligence has a high correlation with both knowledge of English and Chinese."

Dr. Chu also finds that length of time spent in this country has little relation to scholarship, leadership or knowledge of English. Evidently, these are things that depend more on innate ability than on continued experience. Likewise, knowledge of Chinese correlates low with time spent in this country indicating that the forgetting of the mother tongue is a slow process. He concludes that "in order to acquire advanced knowledge in a short period, say, 3 or 4 years, it is almost imperative to send out only those students who have received the A.B. degree in China or who have had a training which will qualify them to enter the graduate schools of America."

Dr. Chu met many difficulties in the statistical treatment of his study, involving as it did the judgment of associates scattered in many universities. He has made two signal contributions along this line. One is a table (p. 13) giving the sigma positions of ranks for a given number of people ranked. This allows comparison of judgments made in different places. For instance, if an individual in University A is ranked first out of 11 people his sigma position is 1.81. If another individual is ranked first out of six people in University B his sigma position is 1.49. The other contribution is that a reliability of 0.90 or over is found from 16 judgments of knowledge of Chinese, 16 judgments of the knowledge of English, 24 judgments of leadership or 20 judgments of scholarship. Criticism of the unreliability of individual judgments has recently been often voiced. *Dr. Chu has given us exact information of the number of independent judgments necessary to achieve any desired reliability from judgment ratings.*

Dr. Chu has produced facts which ought to be of great importance in determining policies regarding the choice and preliminary education of Chinese students coming to this country.

PERCIVAL M. SYMONDS
University of Hawaii.